Routing area
Internet-Draft
Intended status: Standards Track
Expires: September 22, 2016

S. Hegde
Salih. K.A
Juniper Networks
M. Venkatesan
Comcast
R. Callon
A. Atlas
Juniper Networks

March 21, 2016

Virtual multi-instancing for link state protocols
draft-hegde-rtgwg-virtual-multi-instance-01

Abstract

In networks with routers of different capabilities, some routers may not be able to participate fully in supporting new features or handling large databases and the associated flooding. In some cases, these restrictions can cause severe scalability issues for the network in general. This document proposes virtual multi-instances, a generic mechanism for OSPF and IS-IS, that allows groups of routers in specific topologies to have a reduced database and reduces the topology changes that are seen. The virtual multi-instances are specified so that no software or protocol changes are required in the restricted routers. Due to the potential number of virtual multi-instances in a network, the configuration is limited and is not specified per virtual instance.

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].

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 22, 2016.

Copyright Notice

Table of Contents

1. Introduction

A router that participates in OSPF or IS-IS must be capable of handling the entire link state database (LSDB) for the areas or levels that router participates in. In OSPF, this can be mitigated by creating small or stub areas, but such areas must still be configured. In IS-IS, regardless of area address, there can be only a single Level 1. The need to handle the entire LSDB as well as all functionality required in that area or level poses a difficulty for networks that have routers with limited functionality or resources.

In Section 2, the specific problems motivating a solution are discussed. These problems derive from a mixture of operational concerns around configuration, equipment with limited resources, networks with growing numbers of routers, and enhancements in the IGPs that may be needed to support some services but that can't be supported by deployed equipment.

The proposed solution is termed virtual multi-instances because the hub router (termed from a motivating hub-and-spoke topology) is configured to dynamically treat a neighbour's LSP or LSA as belonging to a particular instance, that may be created and deleted on demand. For OSPF, that virtual instance may instead be treated as a virtual area. The hub router automatically creates the virtual instance, distributes a default route into the virtual instance, may advertise specific links into the virtual instance, and redistributes optionally summarized routes learned from that virtual instance. Although the solution does not require any extension to existing protocol standards, the redistribution behaviour should be followed by hub routers for each of the topologies and hence the need for standardization of this solution.

The topologies to which virtual multi-instances can be applied are restricted. In Section 3, the three different types of topologies are described with different behaviour for route redistribution, leaking of hub to hub links into the virtual instance, and ensuring a single hub router LSA/LSP announcement into the virtual instance/ area. The virtual instance or area is distinguished based upon the hub router's and neighbour's Router-ID or system-id or upon the neighbour's specified area-id. An overview of the solution is given in Section 4.

In Section 5, the specific procedures that a hub router must follow to use virtual multi-instances are defined. Because this solution is intended to be low-touch to ease manageability, the suggested

configuration aspects are described in Section 6. In Section 8, the potential security benefits of reducing network visibility and using different instances are briefly discussed.

2. Problem Description

Hub-and-Spoke topologies are increasingly being used at large scale. Due to the scale and to improve routing between spokes, dynamic tunnels between spokes can be established and torn-down on-demand based on traffic flow. Particularly when combined with routers that have limited resources and low-feature implementations of IS-IS or OSPF, these topologies causes real issues in existing networks as described in Section 2.1.

In a hub-and-spoke network, each spoke in the same area unnecessarily learns the link information of the other spokes. This extra information not only grows the size of the LSDB but also causes additional information flooding with associated SPFs. In OSPF, spokes can be separated into different areas but this comes with configuration overhead and can waste IP addresses, since a different IP address is required per interface per area it is used in. In IS-IS, because there is only one L1 domain, the only way to create separated domains is to have separate L1L2 routers for each domain. While [RFC7356] defines different flooding scopes for IS-IS, the changes are not backwards compatible and how the information would be properly processed for basic routing is not defined. In a network, it is rarely feasible to have multiple L1L2 routers in the same geographic area simply to separate the flooding domains.

To provide improved routing between spokes, the ability to establish and tear down dynamic tunnels between spokes on-demand is defined in, for instance, "Auto Discovery VPN Protocol" [I-D.sathyanarayan-ipsecme-advpn]. A huge number of dynamic tunnels can badly impact the scaling of a link-state protocol. At the same time, these on-demand tunnels can’t require configuration overhead to separate them into different areas.

2.1. Issues

2.1.1. Load on Spoke routers

As discussed, containing a hub-and-spoke network inside a single area means that all routers carry the full LSDB for the area. This can overload limited-capability routers or non-router devices that are frequently used as spoke routers. The use of a limited-capability router can thus constrain the size of the area.

High Internet traffic growth requires a high number of link and node updates in metro networks. The number of IP prefixes processed in LS databases increases, causing longer SPF calculations. Though modern routers have high CPUs and better resources for faster SPF calculations, non-router devices typically have limited resources for processing. The size of the LSDB and frequency of SPFs is a problem for non-router devices participating in the routing protocol.

2.1.2. Customer Routers Causing Frequent SPFs

In some cases, service-provider-managed CPEs may participate in the link state routing protocol to advertise their connected and loop-back interfaces for end-to-end connectivity. Power cycles and device failure of CPEs can trigger updates to and SPF calculations on all routers in the domain or area. Isolation of the CPEs from uninvolved routers is desirable.

2.1.3. Mixture of Capabilities between Routers

A metro L1 network supports many different customers and services, but the inclusion of non-router devices (such as cable modem termination systems, video edge devices, voice soft switches, etc.) that participate in the link state protocol may severely limit the ability to provide those different services and abilities.

A non-router device typically just gets its default route from the upstream L1L2 routers for outbound traffic. While that meets the requirements of the non-router device, the inability of such devices to support all IS-IS features (e.g. multi-topology) means that the whole Metro L1 network can’t support those features.

It may not be reasonable or economical to request the implementation of such features on a non-router device that has no need to use them. A solution is required that can support both non-router devices with limited routing protocol features and core network devices with complete routing features. This will allow the Metro L1 network to provide diversified services to different customers.

3. Topology Restrictions

The issues discussed in Section 2 centre on issues around hub-and-spoke topologies. In the simplest case, each spoke is connected to a single hub router as shown in Figure 1. To provide resiliency, a spoke may be connected to two or more hub routers, as shown in Figure 2. Since normal link state routing is performed between the hub and the spoke, the spoke does not need to be a single router, but could be a small connected group of routers operating as an IS-IS

(level 1) or OSPF area as long as only one among the group of routers connects to the hub router.

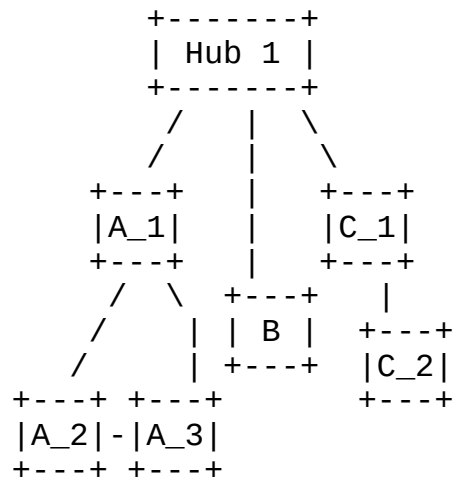

Figure 1: Different spokes connected to a single Hub

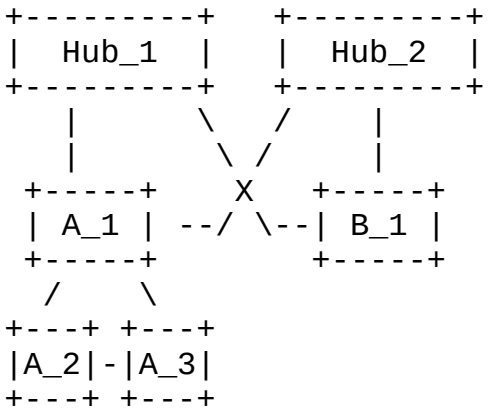

Figure 2: Spokes connected to two Hubs

When there are a huge number of spoke routers, the spokes may be connecting to set of hubs which in turn connect to a hub at the higher level making a hierarchical hub and spoke.It is possible to use virtual multi-instances hierarchically so that a spoke may itself have spokes or rings that have been summarized.

Increased deployments of hub and spoke topologies has lead to improved routing requirements between the spokes. A typical enterprise network with branch offices connecting to head office is usually deployed using IPSEC VPNs. The Figure 4 shows dynamic tunnel topology where A and B are spoke routers and a tunnel is created/ teared-down between them on-demand. The handling of a dynamic tunnel in a virtual instance is slightly different from how a spoke or ring

topology is handled; this is to avoid route redistribution beyond the two ends of the dynamic tunnel.

Another common topology is to have rings that connect to two hub routers, which are themselves directly connected; this is shown in Figure 3; it is possible for additional routers to be connected to the basic ring as shown in ring F. To support ring topologies, the two hub-connecting routers are identified as belonging to the same instance, as described in Section 5 and Section 6. The necessity for this static configuration is what makes it unsuitable for use with dynamic tunnels connecting spokes.

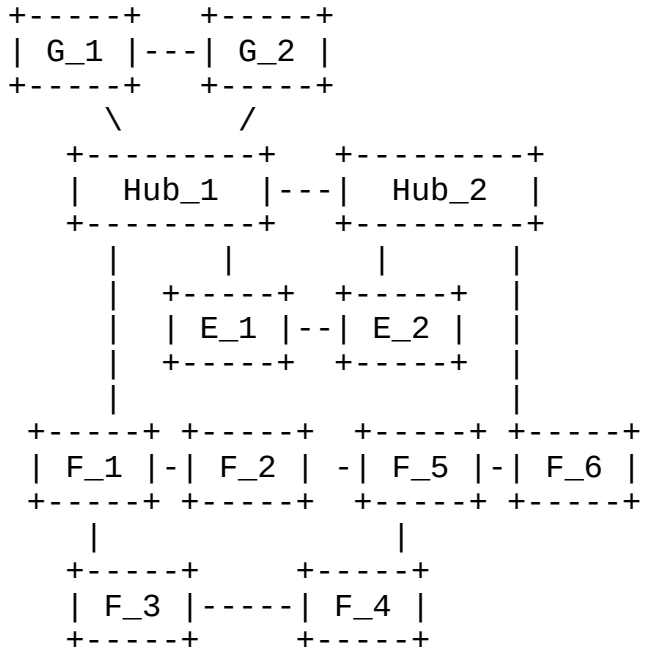

Figure 3: Rings connected to one or two Hubs

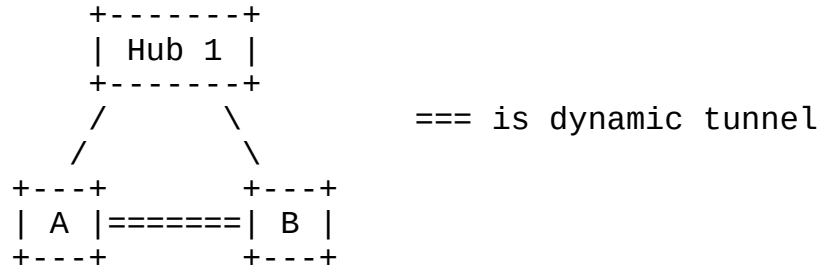

Figure 4: Dynamic tunnel connecting single-node spokes

4. Solution Overview

This document defines virtual multi-instances, which is a mechanism to dynamically create and destroy virtual instances or virtual areas. A similar result can be obtained by creating virtual stub areas in OSPF rather than virtual instances. Whether to create virtual instance or virtual area is an implementation choice.

It is well defined for OSPF and IS-IS how multiple instances can run across a single interface (see [RFC6549] and [RFC6822]) but to support multiple instances in this general case, an instance-id is required in the messages to distinguish which instance is intended. This also requires that all routers in the non-default instance support the extensions. Virtual multi-instances removes the requirement to include the instance-id by both restricting topology and using router-id/system id or area address as keys to distinguish the instances.

By isolating spokes, rings and dynamic tunnels into their own virtual instances, this solution provides isolation for spokes, avoids large LSDBs and, except for handling dynamic tunnels, the need for spoke routers to implement additional features in the IGP. The configuration can be independent of the number of interfaces affected.

4.1. Identification into Virtual Instances or Virtual Areas

There are three different basic types of topologies supported - spoke-based, ring-based and dynamic-tunnel based. A hub router will be configured to know that virtual multi-instance should apply to a set of interfaces and the topology the interfaces correspond to. When an IGP peer is connected via one of those interfaces, the hub router determines the associated instance and, if necessary, creates it. When the last IGP peer disconnects from a virtual instance, the hub router can delete the associated instance. If an IGP peer has a spoke-based or dynamic-tunnel based topology, then the associated virtual instance is identified by the (hub router router-id/system-id, IGP peer router-id/system-id).

In OSPF, it is possible to configure rings as separate stub areas. This requires that all routers in the stub area be configured with the specific and unique area address. In IS-IS, it is not possible to have multiple separate (having separate flooding domain) L1 areas connecting to the same L1/L2 router. For virtual multi-instance to support ring topologies, a router that connects to the hub must be configured with an area address. If multiple routers in the same ring connect to the same hub (routers G_1 and G_2 in Figure 3), then all those and only those routers must be configured with the same

area address. The hub will create a virtual instance or virtual area that is identified by the area address. The hub router does not need to have the area address configured on the set of interfaces to which virtual multi-instance applies. If a single router in a ring connects to a given hub (routers E_1, E_2, F_1, and F_6 in Figure 3, then that router may be configured with a special area address UNIQUE_RING_AREA_ADDRESS (well-known or explicitly configured) and the hub will create a virtual instance or virtual area that is identified by (hub router router-id/system-id, IGP peer router-id/ system-id) but is marked as a ring topology. Virtual instances/areas that are ring topologies will have hub-to-hub links advertised into them.

A router may be connected to a hub via multiple links due to redundancy or to provide sufficient bandwidth. Because a virtual instance is identified by either (hub router router-id/system-id, IGP peer router-id/system-id) or an area address, the multiple IGP adjacencies formed across the parallel links will be in the same instance.

4.2. Route Redistribution

The route redistribution for virtual instances containing a dynamic tunnel is different than that for virtual instances with spoke or ring topologies. For a virtual instance with a dynamic tunnel, only the ends of the dynamic tunnel should learn about the prefixes in the virtual instance. This is to prevent traffic from routing down a spoke and across the dynamic tunnel in order to reach the a destination on the other spoke. A router at the end of a dynamic tunnel

- MUST NOT advertise a default route into
- SHOULD redistribute its own prefixes into
- MAY redistribute non-default prefixes from only its default instance into
- SHOULD NOT redistribute prefixes out of

the associated virtual instance/area.

For spoke and ring topologies, the hub router is responsible for providing a default route into the virtual instance and for redistributing the routes learned from a virtual instance into the default instance. A hub router connected to a spoke or ring topology

- MUST advertise a default route into

- by default, MUST advertise reachability to the addresses that are learned from
- before exporting into the default instance, MAY summarize routes from
- by default, MUST NOT leak routes from the default instance into

the associated virtual instance/area.

Routes from one virtual instance SHOULD not be leaked into each other unless explicitly configured to do so via local policies. By default, routes from default instance MUST NOT be leaked into the virtual instances.

4.3. Avoiding Transit Traffic

Via each virtual instance that is connected to two hubs, a hub router will see a topology to reach the other hub router. However, transit traffic sent via spokes SHOULD be avoided. After the hub router has completed its SPFs in each virtual instance/area as well as any non-virtual instances, the hub router must determine which route is preferred. Routes learned via a non-virtual instance MUST be preferred over routes learned via a virtual instance/area.

4.4. Including Hub-to-Hub Links for Ring LSDBs

Rings that include two hubs usually also need to see the link between the two hubs in their LSDB. This provides redundancy and the possibility of fast-reroute techniques. The link between the hubs is in the default instance. The hub-to-hub links will be advertised by a hub router into all virtual instances/areas that are known to have a ring topology. A hub router can identify other hub routers either by configuration or by using determining other routers with the appropriate node admin tag (see [I-D.ietf-ospf-node-admin-tag] and [I-D.ietf-isis-node-admin-tag]) in the default instance.

4.5. Hierarchical Virtual Multi-Instance

When considering the use of tunnels to connect spokes towards a hub, it is possible for hub-and-spoke topologies to scale extremely high. To reduce the load on particular hubs, it may be useful to consider topologies that include hierarchy so that a spoke router could act as a hub for several remote spokes. Since the spoke router is deliberately unaware that its default instance is being treated as a virtual instance, there are no additional requirements on a router supporting virtual multi-instance.

4.6. Dynamic shortcut Tunnels

As previously discussed (see Section 2), virtual multi-instances need to handle large numbers of dynamic tunnels being created and removed. By way of an example, consider Figure 1 where router A_2 has a dynamic tunnel created to C_2. Router A_2 will create a virtual instance (A_2, C_2) and may redistribute the prefixes associated with C_2, C_1, and Hub_1 into A_2's default instance. Similarly, C_2 will create a virtual instance (C_2, A_2) and may redistribute the prefixes associated with A_1, A_2, A_3, and Hub_1.

Treating a dynamic tunnel as a virtual instance is how dynamic tunnels need to be handled to avoid multiple different LSAs from the same hub router being seen by routers in the connected spokes. Supporting dynamic tunnels does require that router-ends of the dynamic tunnel router support the virtual multi-instance functionality as a hub. There are specific different rules for handling route redistribution (see Section 4.2 for a virtual instance that contains a dynamic tunnel.

In a common topology such as shown in Figure 4, the two spokes each contain a single router A or B and those routers are connected by a dynamic tunnel. In some deployments, it is likely that all connections from router A are sub interfaces across a single interface and that single interface is configured for the "dynamic tunnel topology". In that case, A may treat both the dynamic tunnel to B and the connection to Hub_1 as separate virtual instances and follow the route redistribution rules for the "dynamic tunnel topology" for both. Hub_1 can treat A as being in a "spoke topology" and thus redistribute the needed default route in and redistribute the routes learned from A. This combination will provide the correct behaviour.

5. Hub Router Behavior

5.1. Classification into an Virtual Instance/Area

A received hello, LSupdate or LSP packet needs to be classified as to which instance it belongs to. The following describes how a Hub Router MUST do this classification.

1. Was the LSP/LSA received on an interface configured for virtual multi-instance? If no, select default instance or instance-id in packet and exit.

2. Was the LSP/LSA received on an interface configured for ring topology? If yes, goto (4).

3. Assign to instance or area identified by (hub router-id/system-id, IGP peer source router-id/system-id) and exit.

4. Is the Area ID in the packet the UNIQUE_RING_AREA_ADDRESS? If yes, assign to instance or area identified by (hub router-id/system-id, IGP peer source router-id/system-id) and exit.

5. Assign to instance identified by the Area ID and exit.

New virtual instances/areas SHOULD be created when there is no corresponding instance. With the closing of the last IGP adjacency associated with a virtual instance/area, that virtual instance/area MAY be destroyed.

5.2. Generating Router LSA/LSPs and Default Routes for Virtual Instances

For each virtual instance/area, the Hub Router MUST generate a separate Router LSA/LSP that includes only the links to IGP peers identified as part of that virtual instance/area and, if the virtual instance/area is identified as a ring topology, SHOULD include any direct links from the Hub Router to another Hub Router.

5.3. SPF computations and Route Preference

A separate SPF calculation SHOULD be done for each virtual instance. If the same prefix is learned from a non-virtual instance/area, then its route MUST be preferred over the route via a virtual instance/area.

6. Manageability considerations

Because of the scale for hub-and-spoke topologies, it is difficult to manage per-spoke configuration on the hubs. Therefore, virtual multi-instance does not require per-spoke configuration. The following are the expected configuration aspects.

- A set of interfaces is specified as configured for virtual multi-instance. The type of topology - spoke, ring, or dynamic tunnel - must be specified for the set.

- The set of neighboring hub routers may be specified. This is per hub neighbor configuration. Alternately, node admin tags may be supported and one admin tag configured to indicate what other routers are hub routers.

- A value for the UNIQUE_RING_AREA_ADDRESS may be specified or a well-known default (TBD) may be used.

- A default route to be advertised into virtual instances/areas may be defined.

- A summarization policy for redistributing prefixes may be defined. Ideally, this should apply to a set of virtual instances/areas.

It is expected that virtual multi-instance will be useful to provide a zero-touch hub for IPSEC VPNs where it is highly desirable to have no per spoke configuration on the hub router.

7. Backward compatibility

The mechanism described in the document is fully backward compatible. The mechanism described in this document need to be supported by the hub and the spokes need not support the mechanism unless they need to support dynamic tunnels.

8. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC2328] and [RFC5340].

When a new spoke connects to the hub, it is restricted in terms of visibility into the network. This enhances security in terms of limited exposure to the unauthenticated nodes.Also the ability of a spoke to perturb the entire area is minimized when summarization is done. Per spoke authentication is already available and is expected to work well with virtual multi-instance.

9. IANA Considerations

This document does not currently require any allocations from IANA.

10. Acknowledgements

The authors would like to thank Jeffrey Zhang, Pushpasis Sarkar and Gil Spolidoro for their suggestions and review.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <http://www.rfc-editor.org/info/rfc2328>.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <http://www.rfc-editor.org/info/rfc5340>.

[RFC6549] Lindem, A., Roy, A., and S. Mirtorabi, "OSPFv2 Multi-Instance Extensions", RFC 6549, DOI 10.17487/RFC6549, March 2012, <http://www.rfc-editor.org/info/rfc6549>.

[RFC6822] Previdi, S., Ed., Ginsberg, L., Shand, M., Roy, A., and D. Ward, "IS-IS Multi-Instance", RFC 6822, DOI 10.17487/RFC6822, December 2012, <http://www.rfc-editor.org/info/rfc6822>.

11.2. Informative References

[I-D.ietf-isis-node-admin-tag] Sarkar, P., Gredler, H., Hegde, S., Litkowski, S., Decraene, B., Li, Z., Rodriguez, R., and H. Raghuveer, "Advertising Per-node Admin Tags in IS-IS", draft-ietf-isis-node-admin-tag-08 (work in progress), December 2015.

[I-D.ietf-ospf-node-admin-tag] Hegde, S., Shakir, R., Smirnov, A., Li, Z., and B. Decraene, "Advertising per-node administrative tags in OSPF", draft-ietf-ospf-node-admin-tag-09 (work in progress), November 2015.

[I-D.sathyanarayan-ipsecme-advpn] Sathyanarayan, P., Hanna, S., Melam, N., Nir, Y., Migault, D., and K. Pentikousis, "Auto Discovery VPN Protocol", draft-sathyanarayan-ipsecme-advpn-03 (work in progress), October 2013.

[RFC7356] Ginsberg, L., Previdi, S., and Y. Yang, "IS-IS Flooding Scope Link State PDUs (LSPs)", RFC 7356, DOI 10.17487/RFC7356, September 2014, <http://www.rfc-editor.org/info/rfc7356>.

Authors' Addresses

Shraddha Hegde
Juniper Networks
Embassy Business Park
Bangalore, KA 560093
India

Email: shraddha@juniper.net

Salih K.A
Juniper Networks
Embassy Business Park
Bangalore, KA 560093
India

Email: salih@juniper.net

Mannan Venkatesan
Comcast
1800 Bishops Gate Blvd
Mount Laurel , NJ 08053
USA

Email: mannan_venkatesan@cable.comcast.com

Ross Callon
Juniper Networks
Westford , MA 01886
USA

Email: rcallon@juniper.net

Alia K. Atlas
Juniper Networks

Email: akatlas@juniper.net