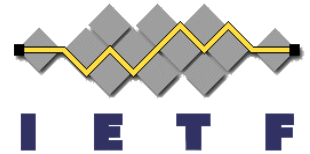


LSVR

(Link State Vector Routing)

Gunter Van de Velde
Victor Kuarsingh



IETF103 Routing Area Open Meeting
8 October 2018
Bangkok, Thailand

DC Routing – What is the fuss about?



- A good motivation overview why DC Routing needs special care and a comparison between various technologies can be discovered in the deck of Tony Li and Chris Martin at NANOG74

“Routing in Dense Topologies What’s All the Fuss?”

https://pc.nanog.org/static/published/meetings/NANOG74/1763/20181003_Martin_Routing_In_Dense_v1.pdf

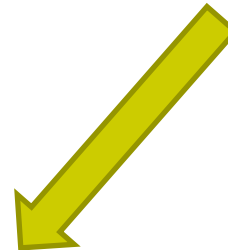
Agenda

- Where are we with LSVR?
 - LSVR WG Chartered schedule
 - LSVR milestone deliverables
- Understanding LSVR
 - Components
 - Motivation for LSVR to use BGP-LS
 - High level technology view
 - LSVR peering models
 - Route decision process (its simple and intuitive)
 - Advantages of LSVR using BGP-SPF
- So what is HOT in LSVR land?
 - Link neighbour liveliness and discovery

Where are we?

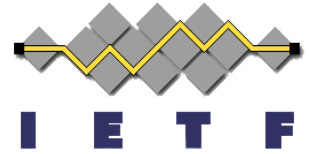
- ~~IETF101 – first stab at the LSVR deliverables~~
- ~~Interim #1~~
- ~~IETF102 (July2018)~~
- ~~Interim #2 (Oct2018)~~
- **IETF103 (November 2018)**
- **Interim #3**
- **IETF104 (March 2019) – Milestone date**

We are here



LSVR Milestones

- March 2019
 - Applicability statement for LSVR in DCs **Adopted**
 - LSV distribution using BGP transport **Adopted**
 - LSVR with standard Dijkstra path selection **Adopted**
- July 2019
 - YANG specification for LSVR



LSVR Components

LSVR using BGP-SPF

Dijkstra SPF

Routing Table Calculation

BGP

Router Link State Distribution

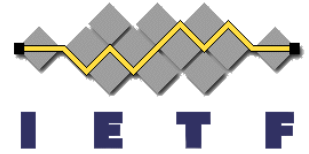
BGP-LS encoding

Router Link State Encoding

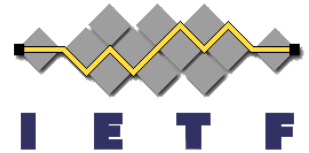
Harvest Link State
LSoE or LLDP

Router Link State Information

Why BGP-based Solution for LSVR?



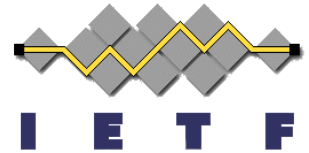
- BGP already widely deployed as sole protocol (see RFC7938) in MSDCs
- Robust and simple implementation
- Wide acceptance – minimal learning
- Reliable transport
- Guaranteed in-order delivery
- Incremental updates
- Incremental updates upon session restart
- No flooding and selective filtering
- Multiple peering models with/without using RR or controllers



High Level LSVR

- Target use-case: MSDC
 - Scale to BGP size massive networks
 - Keep topological view on all BGP speakers (LFA, SRLG, TE extensions, etc)
 - Original target is for underlay routing in MSDC
- Proven BGP-LS encoding is used for LSVs (Link State Vectors)
 - New SAFI proposed for backward compatibility
 - BGP MP capability and BGP-LS node attribute
- Proven re-use of BGP NLRI distribution
 - Route-reflector and controller setup supported
 - Simplified BGP decision process (more rapid)

High Level LSVR



Router Link State Description



= Link State Vector

(neighbor, cost, identity, etc...)

= BGP-LS NLRI Encoding

(new SAFI proposed)

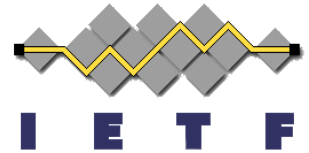
R1

R2

R3

R4

R5



Peering Model

- BGP sessions with Route-Reflector or controller hierarchy
 - Link discovery/liveliness detection outside of BGP-SPF (WIP)
- RR hierarchy can be less than fully connected but must provide resiliency
 - Must not be dependent on SPF for connectivity
- Controller could learn the expected topology through some other means and inject it
 - SPF computation is distributed through
 - Similar approach as :Jupiter Rising: A decade of Clos Topologies and Centralized Control in Google's Datacenter Network)
- BGP SPF Applicability draft covers in detail

BGP Decision Process

Classic BGP Decision Process

Path Vector Algorithm

Phase 1

Calculation of Degree of Preference

Phase 2

Route Selection

Phase 3

Route Dissemination

Simplified BGP Decision Process

LSVR Algorithm

Phase 1&2

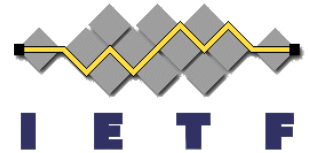
SPF Algorithm

(BGP NLRI Selection & SPF (Route Selection))

Phase 3

NLRI Propagation

(simplified because NLRI is unique per BGP speaker)



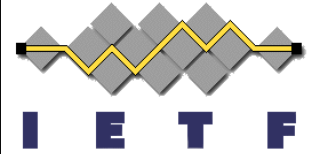
Advantages of BGP-SPF

- Nodes have complete view of topology
 - Ideal when BGP is used as an underlay for other BGP address families
- Only network failures (e.g. link) need be advertised instead of all routes impacted by failure
 - Faster convergence
 - Better scaling
- SPF lends itself better for optimal path selection in Route-reflector and controller topologies

So what is HOT in LSVR?

- Drafts are simple to consume, in good shape and well written
- BGP-SPF draft went successful through OPS-DIR and RTG-DIR review cycle
- Deliverables can still benefit from operator review
- Link Discovery/Liveliness has been discussed for 2 IETFs and Interims
 - It also touch upon IDR playing field requirements

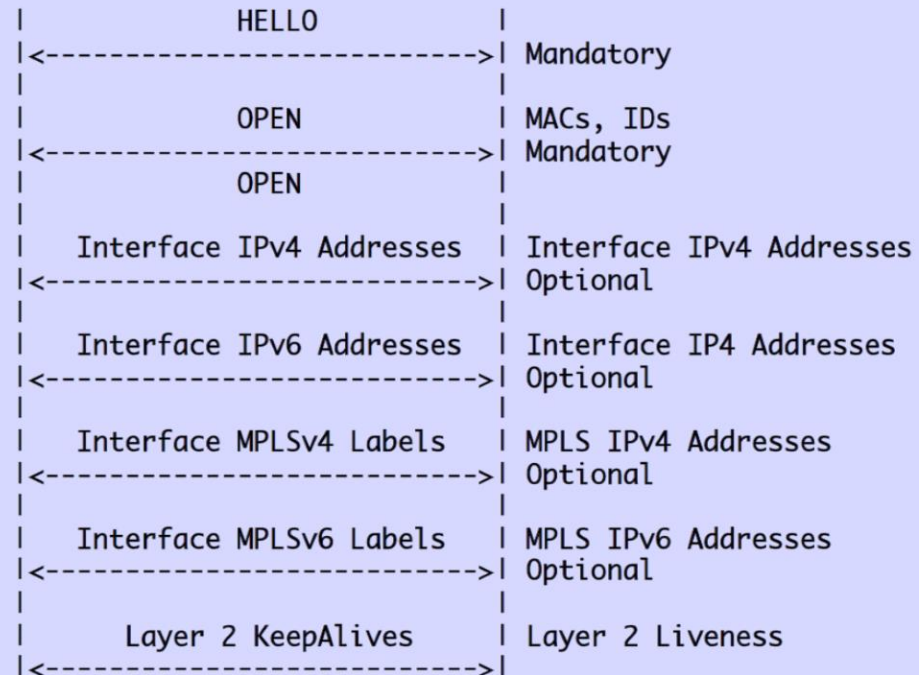
Link State over Ethernet (LSoE)

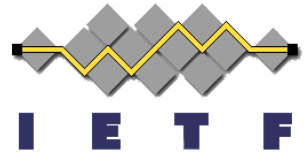


Motivation

- BGP-SPF needs link neighbor discovery, liveness, and addressability
- LLDP is an IEEE protocol, complex, and 'hard' (IPR) to extend past 1500 bytes
- We wanted something simple and saw no real need for the complexities of CLNP, ...
- So we propose a new EtherType with TLVs
- We discuss Ether payloads, not framing

Inter-Link Ether Protocol





THANK YOU!