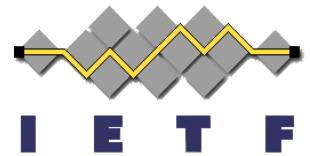


BGP-Based SPF

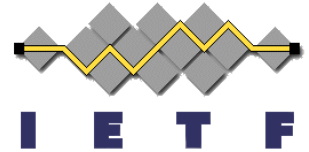
IETF 102 - Montreal

Keyur Patel, Arccus
Acee Lindem, Cisco
Shawn Zandi, LinkedIn
Wim Henderickx, Nokia
Derek Yeung, Arccus
Abhay Roy, Arccus
Venu Venugopal, Cisco



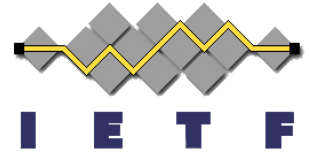
Status Update

- Draft was updated on 05-31-18 prior to interim based on comments
- No draft updates since Interim
- What more is needed in Draft?
 - Main focus of presentation
- Network Operator Interest?
- Implementations



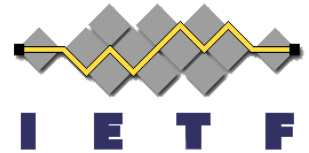
BGP SPF Algorithm (1/3)

- Mark routes corresponding to BGP SPF stale in local RIB
- Start with computing router's Node NLRI and Links as root
 - Add local prefixes to local RIB with metric
 - Optionally, install prefixes corresponding to links
- Add Node NLRI to candidate list based on link cost or simply assure lowest cost is on top (Heap)
 - Only add Node NLRI with corresponding Link NLRI (Bi-directional)



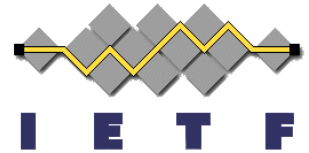
BGP SPF Algorithm (2/3)

- Remove lowest cost Node NLRI off candidate list.
 - Process local prefixes and install if lowest cost maintaining ECMP
 - Optionally install prefixes corresponding to links if lowest cost maintaining ECMP
 - Optionally update routes to node NLRI maintaining ECMP (separate Router RIB or combination)
- Process links corresponding to Node just removed adding nodes that aren't already on the candidate list.



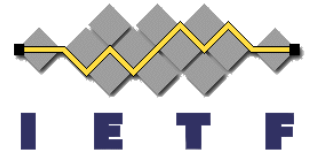
BGP SPF Algorithm (3/3)

- Process links corresponding to Node just removed adding nodes that aren't already on the candidate list
 - Honor bidirectional connectivity check (must be corresponding Link NLRI back to parent node)
 - Maintain ECMP Node NLRI that has equal cost path
- Repeat starting at top of previous slide until the candidate list is empty.
- Determine added, deleted, and changed routes to be installed from BGP Local RIB to Global RIB.



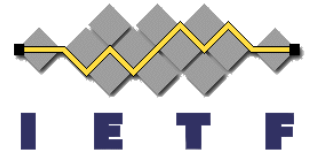
Base SPF versus Strict SPF

- Algorithm 0 - Shortest Path First (SPF) algorithm based on link metric.
 - Algorithm 0 permits any node to overwrite the SPF path with a different path based on its local policy
- Algorithm 1: Strict Shortest Path First (SPF) algorithm based on link metric.
 - Identical to algorithm 0 but requires that all nodes along the path will honor the SPF routing decision irrespective of local policy.
- Only applicable to Segment Routing in IGP today.



Application of BGP Policy

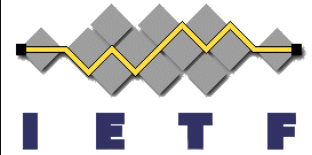
- No inherent concept of areas as in IGP's
 - Not required for data centers
 - Could be added with follow-on draft
- BGP Policy should be provided to roll-your-own area policy at any point. Implementations should provide:
 - Inhibit advertisement of NLRI learned from one neighbor to other neighbors
 - Aggregate prefixes and originate Prefix NLRI based on local RIB routes (Fixed or Maximum Metric)
 - As with current BGP policy, misconfigurations can result in loops and black holes.



BGP-LS SPF Domains

- No inherent concept of areas as in IGP...
- However, do we need separate domains based on neighbors and new NLRI identifier?
 - BGP-LS SPF NLRI would be filtered based on domain identifier
 - One or more domains could be specified for a neighbor
 - New capability necessary or just drop NLRI not matching neighbor domains?
- Redistribution and aggregation would be supported between domains

Decision Process Phase 1 – Determining Best Path (1/2)



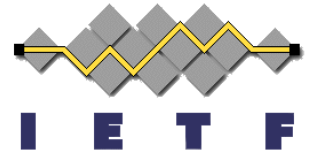
- BGP-LS SPF NLRI originated by single router (unique identifiers)
- Rules for determining best-path
 - Always prefer NLRI from the originator (i.e., Node Identifiers in NLRI match BGP Peer identify)
 - Handles case of BGP SPF Router cold-start complete loss of sequence number state
 - Otherwise, Use sequence number to determine most recent (best-path) NLRI
 - No further criteria necessary since BGP-LS NLRI has single originator

Decision Process Phase 1 – Determining Best Path (2/2)



- Normal BGP BRIB maintenance will keep stale versions of NLRI until they are updated, withdrawn explicitly, or implicitly (BGP session goes down and no graceful restart)
 - Do we want to modify this to converge like IGPs?
 - In other words, do we want to implicitly withdraw stale NLRI or at least mark as stale when better (more recent) NLRI is received?
 - Could really benefit from prototype implementation experience here.

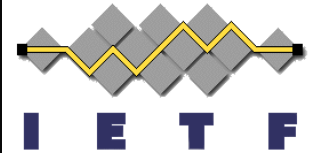
Backup Slides



Changes to Draft version 1

- BGP Peering models section updated
 - Text added to suggest using BFD for link liveness
- BGP SPF SAFI section updated
 - Removed support for VPN SAFI as it is not used in RFC 7752 and the VPN SAFI is applicable to overlay
- Node NLRI section updated for SPF algorithm types
- Added text for relevance of MRAl timer

Changes to Draft version 1 (Cont'd)



- Rearranged the text for BGP Decision process for better readability
- Next_hop manipulation section augmented to cover MP_REACH_ATTRIBUTE nexthop
- IANA section simplified and the new attribute values will be assigned from existing “BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs”
- Finally fixed all the reported Editorial Nits