What's Up in IETF Routing?

Alia K. Atlas

IETF Routing Area Director

akatlas@juniper.net

Outline

- Area-Wide Themes
  - YANG Modeling
  - Encapsulations
  - Centralized Orchestration
- New Working Groups
  - BIER: Bit-Indexed Explicit Replication
  - DetNet: Deterministic Networking
- Improving IETF Work

Yang Models: Participation

- Involved Operators
  - OpenConfig very active
  - High interest in discussions with operators
- Chance to shape and improve common device configuration.
- Many Routing WGs working on YANG models with focused design teams.
- Interesting to look at:
  - OSPF, Traffic-Engineering, BGP

Yang Modeling: Challenges

- Different devices need different sets of models
- Defining full models to be useful for different devices
- Agreeing on common abstractions and functionality to model
- Driving abstractions and model separation from operator perspective, implementations, or protocol?

Classifying Different Models

- Network Service Models (e.g. L3VPN SM)
- Network Element Models
  - Focused towards Configuration
  - Towards learning network (e.g. topology – L2, L3, TE)
  - Towards APIs and dynamic feedback (e.g. RIB)
- How do Models interconnect?
- How to improve reuse of groupings and models?

Yang Modeling: Coordination

- Coordination: rtg-yang-coord@ietf.org
- Routing Yang Architecture Design Team
  - Identify common models and groupings for reuse
  - Describe how models interconnect
  - Raise issues from modeling Routing Area protocols
- Routing Area WG (rtgwg):
  - General forum for routing YANG models without specific WG
    - E.g. Routing-Policy, Key-Chain, RIP, VRRP, etc.
    - Design-Team output:
      Network Device YANG Organizational Model

Outline

- Area-Wide Themes
  - YANG Modeling
  - Encapsulations
  - Centralized Orchestration
- New Working Groups
  - BIER: Bit-Indexed Explicit Replication
  - DetNet: Deterministic Networking
- Improving IETF Work

Lots of Encapsulations

- More Overlay Networks
- Examples: VXLAN-GPE, GUE, GENEVE, NSH, BIER, DetNet??
- Motivations and Differences
  - Transport additional per-packet data
  - How much flexibility?
  - Hardware friendly for what hardware?
  - Avoiding firewalls
  - Support different transports (MPLS, IPv6, UDP) or Transport Independence (for some)
  - Isolated Ecosystems

Encapsulation Considerations

- New encapsulations add HW cost & delay.

- Proprietary encapsulations can be fast & take advantage of specific HW.

- Finding consensus for a single standard in a domain is challenging after deployments.

- Entropy for ECMP
- Packet Size & fragmentation/reassembly
- OAM
- Next Header Indication
- Security & Privacy
- Congestion Considerations
- QoS / CoS
- Header Protection
- Extensibility
- Layering of multiple Encapsulations
- HW-friendly & SW-friendly

Seeking Commonality

- 2015:
  Data-plane Encapsulation Design Team to encourage common solutions where differences aren't needed.

- 2016: Overlay OAM DT to Propose Common/Generic OAM Extensions/ Protocols.

- Complexity challenge: How many different encapsulations for the same purpose are needed?

  – Different Transports: MPLS, IPv6, UDP

  – Different Environments: SW-friendly, HW-

Outline

- Area-Wide Themes
  - YANG Modeling
  - Encapsulations
  - Centralized Orchestration
- New Working Groups
  - BIER: Bit-Indexed Explicit Replication
  - DetNet: Deterministic Networking
- Improving IETF Work

Orchestration & Control-Planes

- Industry movement towards software-defined networks
  - "Centralized Orchestration" (e.g. OpenStack, Contrail, ODL, OVSDB, etc.)
  - PCE and PCEP used as components
- Improved Interaction to Routing
  - Topology learning via BGP-LS
  - BGP Flowspec for traffic direction (many drafts)

Are Models Sufficient?

- Different ecosystems with different control-plane protocols
  - Common need to refer to information for management and monitoring.
  - (proprietary) ability to translate YANG models to other formats
- Do NetConf/RestConf and YANG models suffice? With I2RS extensions and API-like models?
- Model-driven control protocols?
- Continued reuse of existing protocols?

Outline

- Area-Wide Themes
  - YANG Modeling
  - Encapsulations
  - Centralized Orchestration
- New Working Groups
  - BIER: Bit-Indexed Explicit Replication
 - DetNet: Deterministic Networking
- Improving IETF Work

BIER

- First multicast-specific packet encapsulation
- Interesting possible applications (MVPN, Application-specific multicast, etc.)
- Requires new type of forwarding table look-up (*HW changes* or SW implementations)
- Experimental to encourage investigation and work
- Open questions:
  - Does it simplify operations?
  - What can stateless massive-scale multicast enable?
  - Is it useful enough to deploy? What are the challenges to deploy?

BIER Basics

- Each packet carries a BitString, indicating edge routers (within a domain) that need to receive the multicast packet
  - Each edge router has a unique ID, mapping to a bit in the BitString
  - Large domains with potentially long BitString can be handled in various ways of using small BitString
- Each hop looks at the BitString and replicates the packet to its neighbors that are on the shortest paths to the set of edge routers
  - This can be done fairly efficiently
- Removes per-tree/tunnel multicast state in the core

BIFT: Bit index Forwarding Table

A ——— B ——— C ——— D

4 (1000)
3 (0100)

E

1 (0001)
2 (0010)

F
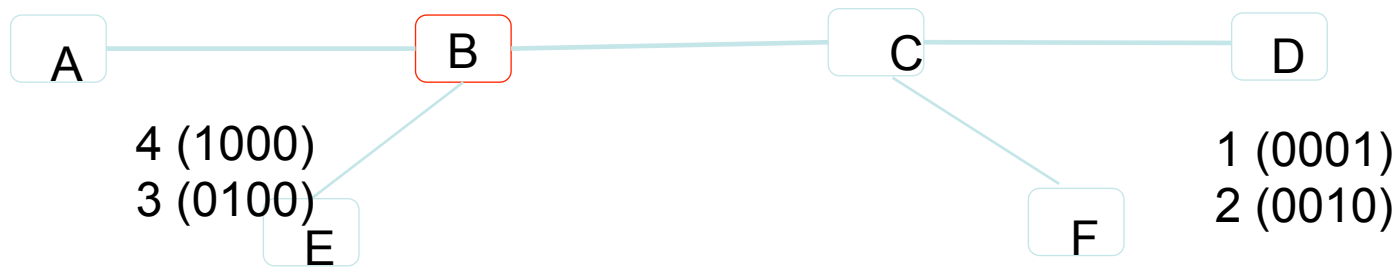
B's BIFT

BFER ID	F-BM	BFR-NBR
1 (0001) D	0011	C
2 (0010) F	0011	C
3 (0100) E	0100	E
4 (1000) A	1000	A

Route Lookup & Packet forwarding

- Send 1 copy to a BFR-NBR that is on the shortest path to a subset of the BFERs that needs to receive the packet
- Start with the lowest set bit in the packet's BitString – use that bit's index to look up the BIFT
  - The row identifies the BFR-NBR (to send a copy to), and a F-BM that indicates all BFERs that are reachable by the BFR-NBR
    - Set the copy's BitString to (packet's BitString & F-BM)
    - Change the original packet's BitString to (packet's BitString & ~F-BM)
  - Repeat the procedure to send another copy to another NBR
- Repeat above N times, where N is the number of neighbors that need to forward the packet to all BFERs that need to receive the packet

Example

BFER ID	F-BM	BFR-NBR
1 (0001) D	0011	C
2 (0010) F	0011	C
3 (0100) E	0100	E
4 (1000) A	1000	A

- Incoming packet with BitString 0111
  - Need to reach E,F,D
- Lowest set bit is the 1st (right most) so use index one to look up BIFT
- Row 1 (index 1) has F-BM 0011 and BFR-NBR C
  - Send a copy to C, with BitString 0011 (to reach F,D)
    - 0011 == 0111 & 0011   (BitString & F-BM)
    - Notice that this takes care two bits (all those BFERs to be reached via C)
  - Change packet's BitString to 0100
    - 0100 == 0111 & 1100   (BitString & ~F-BM)
- Now the lowest set bit is the 3rd so use index 3 to repeat the above
  - Send a copy to E, with BitString 0100

Deterministic Networking (DetNet)

- Grew from 1EEE802.1 Time-Sensitive Networking (TSN) Task Group as needs moved from pure Layer-2 to mixed layer-2/layer-3

- Use-cases from Industrial Automation, Audio/Video, In-Vehicle or Avionics Networks.

- Restricted to a single Administrative Control/Closed Group.

- Route deterministic flows across network with controlled packet latency and loss.

  - Centrally orchestrated or Distributed Path Setup

  - Unicast and Multicast Flows

  - Can be >50% capacity

  - Requires pre-reserved resources (buffers, etc.)

DetNet: Investigating Technologies

– Could identify flows via an MPLS label (G-MPLS for 6Tisch).

– Interest in IPv6 as well; possibly use DSCP to identify the packet-track.

– For the Path Computation Element (PCE), may need to share more device-specific topology and resources information.

– Preliminary architecture gives 3 techniques:

- Zero congestion loss via assigned resources

- Pinned-down Paths or Trees

- Packet Replication and Deletion to handle a failure

Outline

- Area-Wide Themes
  - YANG Modeling
  - Encapsulations
  - Centralized Orchestration
- New Working Groups
  - BIER: Bit-Indexed Explicit Replication
 - DetNet: Deterministic Networking
- Improving IETF Work

Improving IETF: Speed

- It can feel slow (stability has a cost)
  - Volunteers with competing priorities
  - Consensus means compromise & discussion (RFC 7282)
  - Work moves when **operators really want it**. (e.g. BFD over LAG)

- Working on Improvements
  - More WG conference calls (virtual interims) to keep momentum.
  - Monthly Training & discussion with WG Chairs
  - Move drafts through process faster

NO RESOURCES
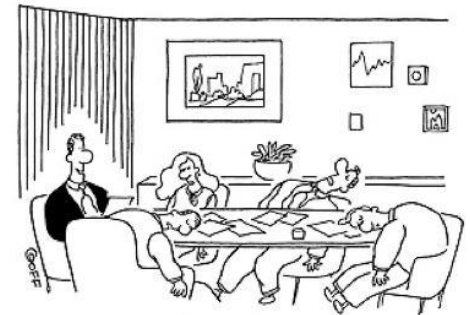
"at last we have reached a consensus!"

Improving IETF: Why Get Involved?

- **Why?** Because you understand your networks and needs better than your vendors.

- **Why?** Have an impact and make the Internet work better.

- **Why?** Find problems with the technology before it hurts your network.

- **Why?** Because operators serve as a forcing function on IETF – what's needed, useful and deployable?

How to Get Involved?

- Review an interesting Working Group draft and respond back to the authors & WG.

- Join an interesting/relevant Working Group mailing list and provide perspective.

- Find your social connections to those already active and ask.

- You don't need to spend lots of time.

https://datatracker.ietf.org/wg/#rtg

How to Get Involved Specifically?

- Comments to IDR: Should BGP FlowSpec be based on the associated YANG models so the same functionality can be done via RestConf/NetConf as via BGP?

- Comments to NVO3: Preferred data-plane encapsulation and reasons (VXLAN-GPE, GUE, GENEVE); feedback on OAM.

- Comments to RTGWG: IPv6 Multihoming in Small Enterprise – what are requirements and trade-offs BCP38?

Q & A