IETF Routing Area Update

A Report after IETF 95 (Buenos Aires)

Alvaro Retana (aretana@cisco.com)
Distinguished Engineer, Cisco Services
Routing Area Director, IETF

No one is in charge, anyone can contribute and everyone can benefit.

IETF Organization: Areas

General Area (gen)	• ...activities focused on supporting, updating and maintaining the IETF standards development process.
Security (sec)	• ...focused on security protocols...services: integrity, authentication, non-repudiation, confidentiality, and access control...key management is also vital.
Applications and Real Time (art)	•Protocols for delay-sensitive communications, and building blocks to be used across a wide variety of applications.
Operations & Management (ops)	• Network Management, AAA, and various operational issues facing the Internet such as DNS, IPv6, operational security and Routing operations.
Transport Services (tsv)	• ...works on mechanisms related to end-to-end data transport...
Routing (rtg)	• ...responsible for ensuring continuous operation of the Internet routing system...
Internet (int)	• ...IP layer (both IPv4 and IPv6), DNS, mobility, VPNs and pseudowires..., and various link layer technologies.

IETF Meetings

IETF 95
April 3-8, 2016
Buenos Aires, Argentina

Recent Meetings

- **IETF 92**
  - March 22-27, 2015
  - Dallas, TX, USA
- **IETF 93**
  - July 19-24, 2015
  - Prague, Czech Republic
- **IETF 94**
  - November 1-6, 2015
  - Yokohama, Japan

Upcoming Meetings

- **IETF 96**
  - July 17-22, 2016
  - Berlin, Germany
- **IETF 97**
  - November 13-18, 2016
  - Seoul, South Korea
- **IETF 98**
  - March 26-31, 2017
  - Chicago, IL, USA

Hot Topics at IETF 95

- Infrastructure Resiliency and Security
  - DNS and Routing
- Scalability and Performance
- IPv6
- YANG Modeling
- IoT
- Trust, Identity, and Privacy

Plenary (Wednesday 1740)

IETF-wide and Architectural Issues

ISOC@IETF Briefing Panel

Public Policy and Internet Technology Development

(Tuesday 1245)

https://datatracker.ietf.org/meeting/95/materials/

http://www.internetsociety.org/rough-guide-ietf95

Routing Area (RTG)

- “...responsible for ensuring continuous operation of the Internet routing system by maintaining the scalability and stability characteristics of the existing routing protocols, as well as developing new protocols, extensions, and bug fixes in a timely manner.”

- 24 WGs (soon to be 25!)
- 3 Area Directors

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

BABEL BoF

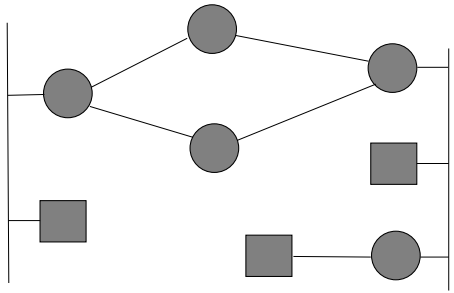
- Proposed Charter
  - *The charter focuses on what is needed to move RFCs 6126 and RFC 7557, combined, to Proposed Standard*
  - *It also recognizes that we need to include security and manageability in such a move*

https://datatracker.ietf.org/wg/babel/charter/

Hybrid networks

Successful deployment 1/4

Babel works well in classical, prefix based networks (supports aggregation, filtering, etc.).
Babel works well in pure mesh networks (non-transitive and unstable links).

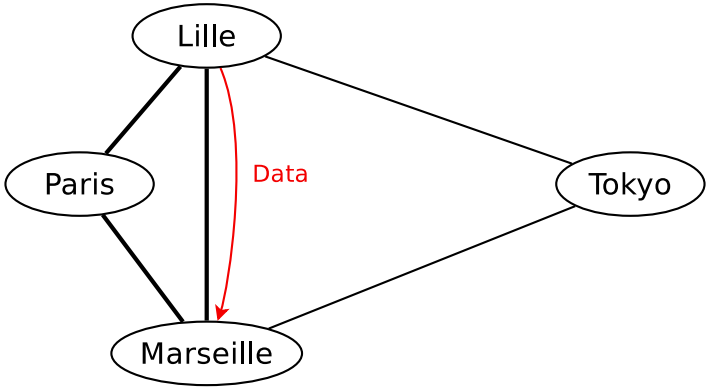


Babel works well in hybrid networks, networks with prefix based parts interconnected through meshy bits.

Global-scale overlay networks

Successful deployment 2/4

The RTT-based routing extension enables non-pessimal routing in global-scale overlay networks:



RTT-based routing may cause persistent oscillations, but Babel remains robust even in the presence of oscillations.

9/14

Source-specific routing

Successful deployment 3/4

The source-specific extension to Babel gives:

- full support for source-specific routing (SADR);
- interoperability with plain, unextended Babel.

Babel is useful wherever source-specific routing is needed.

Small, simple networks

Successful deployment 4/4

Babel is a small, simple protocol and requires no configuration in simple cases.

It is often used in trivial networks: a useful RIP replacement.

Pure mesh networks

Potential deployment 1/1

Babel has been repeatedly shown to be competitive with dedicated mesh routing protocols:

- better on some tests;
- worse on others.

However, standardised, well implemented protocols for mesh networks exist:

- OLSR-ETX;
- OLSRv2 with the DAT metric;
- . . .

This particular niche is already populated.

Large, stable networks

Non-recommended deployment 1/1

There exist protocols that are finely tuned for large, wired networks:

- OSPF;
- IS-IS;
- EIGRP.

Babel relies on periodic route announcements, and will never be competitive with protocols that only send deltas.

13/14

Routing Area (rtg)

- BGP Enabled Services (bess)
- Bidirectional Forwarding Detection (bfd)
- Bit Indexed Explicit Replication (bier)
- Common Control and Measurement Plane (ccamp)
- Deterministic Networking (detnet)
- Interface to the Routing System (i2rs)
- Inter-Domain Routing (idr)
- IS-IS for IP Internets (isis)
- Layer Two Tunneling Protocol Extensions (l2tpext)
- Locator/ID Separation Protocol (lisp)
- Mobile Ad-hoc Networks (manet)
- Multiprotocol Label Switching (mpls)
- Network Virtualization Overlays (nvo3)
- Open Shortest Path First IGP (ospf)
- Pseudowire And LDP-enabled Services (pals)
- Path Computation Element (pce)
- Protocol Independent Multicast (pim)
- Routing Over Low power and Lossy networks (roll)
- Routing Area Working Group (rtgwg)
- Service Function Chaining (sfc)
- Secure Inter-Domain Routing (sidr)
- Source Packet Routing in Networking (spring)
- Traffic Engineering Architecture and Signaling (teas)
- Transparent Interconnection of Lots of Links (trill)

Routing Area (rtg) – IP Routing

- BGP Enabled Services (bess)
- Bidirectional Forwarding Detection (bfd)
- **Bit Indexed Explicit Replication (bier)**
- Common Control and Measurement Plane (ccamp)
- Deterministic Networking (detnet)
- Interface to the Routing System (i2rs)
- **Inter-Domain Routing (idr)**
- **IS-IS for IP Internets (isis)**
- Layer Two Tunneling Protocol Extensions (l2tpext)
- Locator/ID Separation Protocol (lisp)
- Mobile Ad-hoc Networks (manet)
- Multiprotocol Label Switching (mpls)
- Network Virtualization Overlays (nvo3)
- **Open Shortest Path First IGP (ospf)**
- Pseudowire And LDP-enabled Services (pals)
- Path Computation Element (pce)
- **Protocol Independent Multicast (pim)**
- Routing Over Low power and Lossy networks (roll)
- Routing Area Working Group (rtgwg)
- Service Function Chaining (sfc)
- **Secure Inter-Domain Routing (sidr)**
- Source Packet Routing in Networking (spring)
- Traffic Engineering Architecture and Signaling (teas)
- **Transparent Interconnection of Lots of Links (trill)**

Routing Area (rtg) - MPLS

- **BGP Enabled Services (bess)**
- Bidirectional Forwarding Detection (bfd)
- Bit Indexed Explicit Replication (bier)
- **Common Control and Measurement Plane (ccamp)**
- Deterministic Networking (detnet)
- Interface to the Routing System (i2rs)
- Inter-Domain Routing (idr)
- IS-IS for IP Internets (isis)
- Layer Two Tunneling Protocol Extensions (l2tpext)
- Locator/ID Separation Protocol (lisp)
- Mobile Ad-hoc Networks (manet)
- **Multiprotocol Label Switching (mpls)**
- Network Virtualization Overlays (nvo3)
- Open Shortest Path First IGP (ospf)
- **Pseudowire And LDP-enabled Services (pals)**
- Path Computation Element (pce)
- Protocol Independent Multicast (pim)
- Routing Over Low power and Lossy networks (roll)
- Routing Area Working Group (rtgwg)
- Service Function Chaining (sfc)
- Secure Inter-Domain Routing (sidr)
- Source Packet Routing in Networking (spring)
- **Traffic Engineering Architecture and Signaling (teas)**
- Transparent Interconnection of Lots of Links (trill)

Routing Area (rtg) – SDN/Overlays

- BGP Enabled Services (bess)
- Bidirectional Forwarding Detection (bfd)
- Bit Indexed Explicit Replication (bier)
- Common Control and Measurement Plane (ccamp)
- Deterministic Networking (detnet)
- **Interface to the Routing System (i2rs)**
- Inter-Domain Routing (idr)
- IS-IS for IP Internets (isis)
- **Layer Two Tunneling Protocol Extensions (l2tpext)**
- **Locator/ID Separation Protocol (lisp)**
- Mobile Ad-hoc Networks (manet)
- Multiprotocol Label Switching (mpls)
- **Network Virtualization Overlays (nvo3)**
- Open Shortest Path First IGP (ospf)
- Pseudowire And LDP-enabled Services (pals)
- **Path Computation Element (pce)**
- Protocol Independent Multicast (pim)
- Routing Over Low power and Lossy networks (roll)
- Routing Area Working Group (rtgwg)
- **Service Function Chaining (sfc)**
- Secure Inter-Domain Routing (sidr)
- **Source Packet Routing in Networking (spring)**
- Traffic Engineering Architecture and Signaling (teas)
- Transparent Interconnection of Lots of Links (trill)

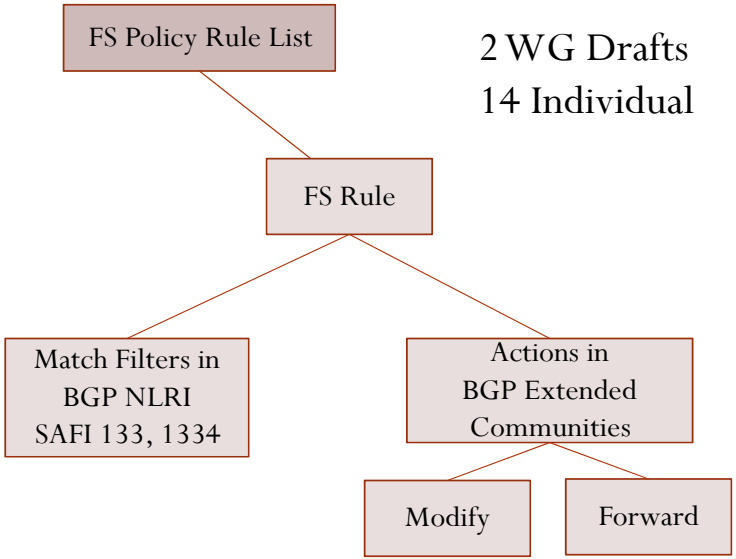
Routing Area (rtg) – Mobility / IoT

- BGP Enabled Services (bess)
- Bidirectional Forwarding Detection (bfd)
- Bit Indexed Explicit Replication (bier)
- Common Control and Measurement Plane (ccamp)
- Deterministic Networking (detnet)
- Interface to the Routing System (i2rs)
- Inter-Domain Routing (idr)
- IS-IS for IP Internets (isis)
- Layer Two Tunneling Protocol Extensions (l2tpext)
- Locator/ID Separation Protocol (lisp)
- **Mobile Ad-hoc Networks (manet)**
- Multiprotocol Label Switching (mpls)
- Network Virtualization Overlays (nvo3)
- Open Shortest Path First IGP (ospf)
- Pseudowire And LDP-enabled Services (pals)
- Path Computation Element (pce)
- Protocol Independent Multicast (pim)
- **Routing Over Low power and Lossy networks (roll)**
- Routing Area Working Group (rtgwg)
- Service Function Chaining (sfc)
- Secure Inter-Domain Routing (sidr)
- Source Packet Routing in Networking (spring)
- Traffic Engineering Architecture and Signaling (teas)
- Transparent Interconnection of Lots of Links (trill)

Routing Area (rtg) - Other

- BGP Enabled Services (bess)
- **Bidirectional Forwarding Detection (bfd)**
- Bit Indexed Explicit Replication (bier)
- Common Control and Measurement Plane (ccamp)
- **Deterministic Networking (detnet)**
- Interface to the Routing System (i2rs)
- Inter-Domain Routing (idr)
- IS-IS for IP Internets (isis)
- Layer Two Tunneling Protocol Extensions (l2tpext)
- Locator/ID Separation Protocol (lisp)
- Mobile Ad-hoc Networks (manet)
- Multiprotocol Label Switching (mpls)
- Network Virtualisation Overlays (nvo3)
- Open Shortest Path First IGP (ospf)
- Pseudowire And LDP-enabled Services (pals)
- Path Computation Element (pce)
- Protocol Independent Multicast (pim)
- Routing Over Low power and Lossy networks (roll)
- **Routing Area Working Group (rtgwg)**
- Service Function Chaining (sfc)
- Secure Inter-Domain Routing (sidr)
- Source Packet Routing in Networking (spring)
- Traffic Engineering Architecture and Signalling (teas)
- Transparent Interconnection of Lots of Links (trill)

BGP Flow Specification RFC5575++



2 WG Drafts
14 Individual

Use Case for

- Option 1: Minimal Flow Specification
  - Use Case: Prevent DDoS
  - How: Expands Filters + Actions with Defined order
- Option 2:
  - Use Case: SDN/NFV central controller for paths or segments
  - Why BGP: Peer distribution of some filters from a certain .
  - How: New NLRI + Wide Community
  - Chair: Call for input from providers
- Not I2RS vs. BGP – 2 tools

Order is needed

Precedence needed within BGP Flow Specification

- For filtering – Currently all
  - For ordering policies: use NLRI preference and administrative distance
  - For ordering filters – by Flow Specification type and precedence
- For action
  - No order currently, need to add order

Common work

- Common
  - Decided if need ROA or draft-ietf-idr-bgp-flowspec-oid-03
  - Define default precedence ordering for filters + Actions
  - Define precedence between BGP-FS and other packet filters (E.g. I2RS FB-RIB)
  - Define conflict resolution between actions
- Choose drafts per Option

BGP FS Filters types for RFC/WG documents

- RFC 5575 types/v6-draft
  1. Destination prefix
  2. Source prefix
  3. IPv4 protocol / IPv6 Next header
  4. Port (source or destination)
  5. Source port
  6. Destination port
  7. ICMP Type
  8. ICMP Code
  9. TCP Flags
  10. Packet length
  11. Traffic Class
  12. IPv4 Fragment
  13. IPv6 Flow ID
- L2VPN types
  14. Ethernet type
  15. Source MAC
  16. Destination MAC
  17. DSAP in LLC
  18. SSAP in LLC
  19. Control fields in LLC
  20. SNAP
  21. VLAN ID
  22. VLAN COS
  23. Inner VLAN ID
  24. Inner VLAN COS

Flow Specification Actions

Approved Actions

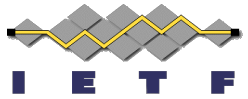
(RFC 5575 & RFC 7674)

- Traffic rate in bytes (0x8006)
- Traffic Action (0x8007) with S(sample) T (terminal) flags
- Redirect to IP VPN via Route Target
  - RD 2 octet AS, 4 byte value (0x8008)
  - RD 4 octet IP, 2 byte value (0x8108),
  - RD 4 octet AS, 2 byte value (0x8208)

Proposed Actions

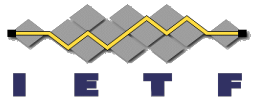
- (FA1) Traffic Rate in packets
- (FA2) Traffic Action with "R" for refer to more policy in BGP Attribute
- (FA3) Redirect to Tunnel
- (FA4) VLAN Action
- (FA5) TPID action
- (FA6) MPLS label action (push, pop, swap)
- (FA7) change validation to ROA or bgpsec-protocol
- (FA8a) interface set
- (FA8b) ACL+BGP FS

This talk triggered by operator request in v6ops



- A number of operators opined at IETF 94 that,
  - while PI multihoming is common and works well,
  - PA is difficult for enterprise to deploy without egress routing
- Those few networks using it resort to operational means such as
  - Flash renumbering
  - Using one ISP's prefix in one place and another ISP's in another
  - Forcing all traffic through a single egress router

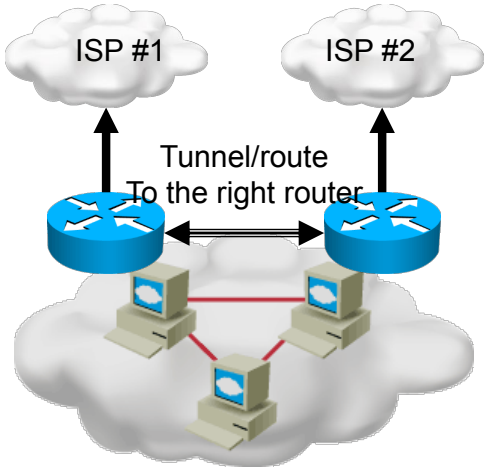
Egress Routing impetus



- IETF generally recommends* use of provider-allocated prefixes in generalized multihoming for smaller networks
  - PI obviously works and is used by larger networks that use BGP and have AS numbers
  - The point is to minimize impact on the global route table by enabling ISPs to aggregate smaller multihomed customers into their own prefix
- Issue:
  - BCP 38 encourages ISPs to drop customer traffic that uses addresses they don't know the customer to be using

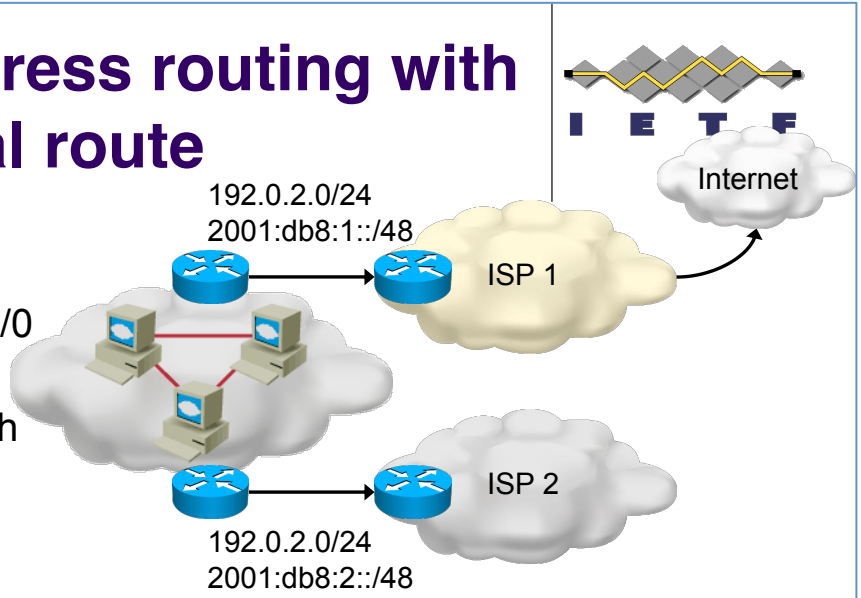
History

- This came to a head in the IETF in 2014, when v6ops WG Chair asked me to write up a solution
  - RFC 3704
- Concept:
  - Destination route within a network
  - At the egress, wonder what source prefix is in use
    - If the correct one for upstream, send upstream
    - Else, re-route to the correct egress router
- My question:
  - Why not route it to the right router in the first place?



Second use case: egress routing with a specialized external route

- Multiple ISPs
  - ISP 1: routing from a prefix to ::/0 (default route)
  - ISP 2: Specialized service (such as NTT BFLETS)
- Specialized ISP offers a destination route to its prefix, and requires network (home) to use its PA prefix when accessing it.



- Yes, you could use destination routing and let hosts learn which source address actually works. If they actually learn.

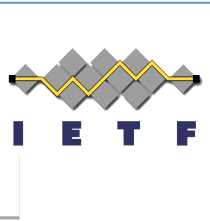
Routing Area Yang Architecture Design Team Update

Members: Acee Lindem, Anees Shaikh, Christian Hopps, Dean Bogdanovic, Lou Berger, Qin Wu, Rob Shakir, Stephane Litkowski, Yan Gang

Wiki: http://trac.tools.ietf.org/area/rtg/trac/wiki/RtgYangArchDT

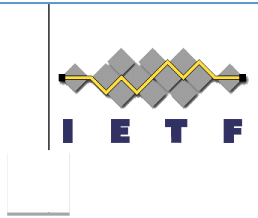
Repo: https://github.com/ietf-rtg-area-yang-arch-dt/

DT current “work” topics



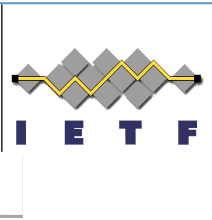
1. Meta-Model:
   YANG Device Model Structure
2. OpState:
   YANG Relationship of Config and Operational State (and intended)
3. Conventions

Status: Meta-Model



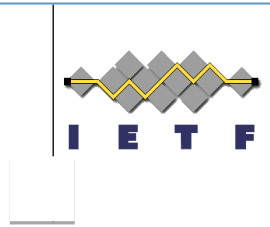
- Significant progress from last meeting!
- Identified need for “schema mount”
  - To simplify organization (more on this later)
    - And remove instance/LNE from all models!
  - NETMOD interim held, NETMOD took action to provide solution
- Published draft-rtgyangdt-rtgwg-device-model-03
  - Assuming schema mount
- Next steps
  - Track schema mount solution development
  - Socialize solution
    - Mostly sync’ed with draft-ietf-netmod-routing-cfg

Status: OpState



- Requirements documented and generally understood
  - Telemetry/streaming use case not called out in netmod req document, but understood by those working solution
- NETMOD solutions authors working on consolidated / unified approach
- Lack of standard OpState solution is blocking other work
  - Options available today:
  1. Ignore OpState – assume solution won't require model changes
  2. Manually add to every model – current OpenConfig approach
- In holding pattern – next DT steps:
  - Track solution discussion in netmod
  - Once there is a solution, sanity check, update drafts as needed

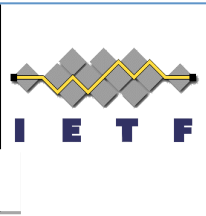
Status: Conventions



Objectives (from AD):

- Provide YANG structure conventions for area
  - E.g., containers within groupings, lists within containers, etc.
- Provide guidance to routing area protocol WGs on:
  - Process for modifying existing models
  - What to do, i.e., not forget about, WRT YANG when defining new protocol extensions
    - TBD, e.g., a new “YANG considerations section”…

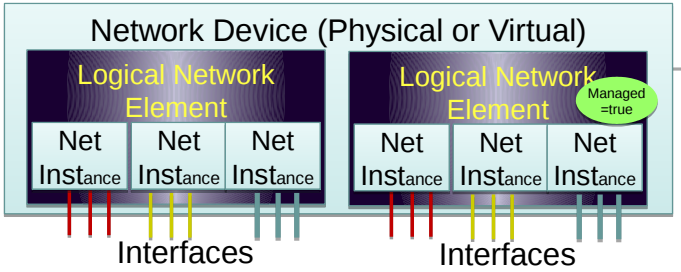
Defined Models



1. module: network-device
   - Overall structure for any network device type
     - From small router to Carrier Class
     - Covers relations amongst models –
       Not to be implemented directly
2. module: logical-network-element
   - Separates management/resource domains
     - Commonly called logical system or router, and virtual switch, chassis, or fabric, virtual device contexts, contexts
3. module: network-instance
   - Separates routing or switching domain
     - e.g., VRF or VSI

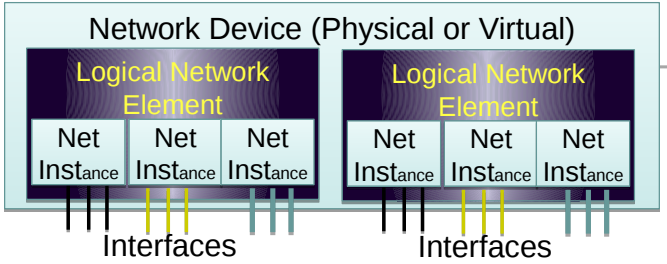
- Will eventually be broken into three documents

Logical Network Elements



- Separate management sub-domains
  - Sub-domains can be managed independently and by a top level manager (managed=true)
- Differs from multiple logical devices and VMs
  - Where top level management of subdomains not supported

Network Instances



- Separate routing / switching domains
- Can represent of an RFC 4364 VRF or a Layer 2 Virtual Switch Instance (VSI) or a bridge/router (i.e., both)
- General virtualized instance implying a separate L2, L3, or L2/L3 context.
  - For L3, this implies a unique IPv4/IPv6 address space.

Summary

- Routing Protocols are mature and stable, but are also entering a new era of increased, dynamic coverage.
  - Convergence, Availability, Scalability and Security are still front and center...
- More than 200 routing-related work items are being considered in the Routing Area (and beyond)
  - New requirements are coming from a diverse set of sources: from the Internet of Things, traditional SP and Enterprise networks, to SDN and beyond.

Get Involved!

TOMORROW starts here.