I2RS Interium WG Meeting: 4/22/2013 - 4/23/2013

Co-chairs: Alia Atlas, Ed Crabbe

9:30 - Getting started

Introduction - Ed Crabbe

note well noted

intro to plan to spend most of day in small groups

few updates to drafts since IETF86. Ed had hoped drafts would get updated as per normal IETF.

Few presos this morning:
- Alia will discuss output we're expecting from small groups.
- Sue will present policy use case examples.
- Ed will present on topology models
- Jan will present addt'l topology use case models

Going to use Google Groups for small group break-out meetings.

Also google tools (for now) for file sharing

After lunch, present feedback from morning small WG sessions.

Day 2: spent almost entirely in small groups. Come back and present what you've put together.

Charter Review
Review of WG milestones: aggressive deadlines with problem statement & high-level arch. due by summer. High Level Arch. & Problem Statement due 7/2013. Use cases due by August 2013.

chartered for problem statement, use cases, architecture etc. need to get that done so we can do the interesting work (protocols etc.)

Small Group Structure: groups are assign a topic, Google Hangout and participation spreadsheet. ~1/3 of attendees are remote, so its intended to help get involvement from remote participants.

topics assigned to groups and hangouts pre-created. structure is intended to help remote participants - clearly easy for local people to cooperate.

goals for small groups: short term milestones - produce actionable plans for drafts etc. When looking at use

case try to determine specific new things required - and what is addressed by existing protocols. good to explictly call the overlap out.

Joel: concern about Arch. Requirements. Don't want people to rathole in Arch. Reqmt's; but, rather to call out what are Arch. Reqmt's for I2RS, where an existing protocol may already exist, etc.

Example of Small Group Work: DDoS protection
- method to describe filters, collect stats, mirror traffic, route traffic based on match, push state updates to devices
- Why are these (existing protocols) not sufficient? What addt'l functionality is required?

Nabil: Question on use case overlap: service chaining vs. service provisioning
- Ex. of Service Chaining - linking a service into a tunnel, etc.;
- Ex. of Service Provisioning is more about realizing a new service at edge of the network.

Ed doesn't want to constrain us to these topics. if you have other use-cases talk to Ed and he'll set up the hangout.

Alia - we need structure. let's see how meet today goes and we'll re-plan for tomorrow.

Q - if 30 people here than 10 groups == 3 per group. Alia - that's ok. V fluid - just trying to put enough structure in to get some output.

Ed - 4/5 people is optimal and should get that given 55 registered to be here.

Himanshu - might be a good idea to combine some groups as some may be too small. 6 BGP groups for example.

Alia - happy to put all BGP sheets together and then BGP people can self-negotiate.

9:55am. Alia -
Example Template of I2RS WG preso
1st, 2nd & 3rd Small Group Outline
1st mtg: general description, scope & scale
- scope & scale: what ypes of areas of network would find this useful; what set of network elements are involved, (two ends of network); how much and how frequent are operations, (read/write/events)?
- what's different for I2RS? how is the use-case currently solved? What would make I2RS better/worse?

2nd mtg: flesh out details of use case; put it in context; describe data and events needed
- put use case in context of network application and the network element
- describe the data and events needed
- detailed use-case: words & pictures to show responsiveness and feedback loop.

- what events are needed; do any events trigger a local action (local to the box?)
- what data is needed? what info is pulled or pushed from the NE? how frequently does this occur?

3rd meeting: reqmt's of use-case on I2RS; commit to WG draft; suggest go/no-go on use case for I2RS

one small group meet is "taxonomy of network applications". different ideas apply to different use-cases. Trying to name that space rather than restrict it. Controller? Network-wide scope? Narrow network service talking to 2 end-points to set up a service. Does controller create state and go away or does it persist? we lack terminology to discuss it.

2nd group meetings on use-cases we can then use that taxonomy? "this is the type of network application" or "these are the attributes the app is expected to have".

Himanshu - it's also about related information. e.g. if you install a static route on box X pointing to Y and vice-versa then you've created a loop. What about conflicting configurations? How do you resolve that, paritcularly across multiple NE's?

Alia: We're not trying to take away your ability to create conflicting configs. Ops & Apps will need to be careful with how they configure the network. No diff. than today.

Ed: Anytime you have apps that are writing distributing state to the network, then there is going to have to be a "global broker" that is going to arbitrate state.

Requirements for 3rd Meeting:

Connect to I2RS Functionality
- Multi-threaded control; precedence per client or per operation
- Types of operations needed
- Event subscription/notification requirements
- Atomic Operations sufficient? Are longer transactions needed?
- Rollback by client sufficient?

How specific does authentication/authorization need to be, and why?

Opinion on use-case for I2RS?
- Is it a good fit for I2RS? Why or why not?
- NEEDS TO BE WRITTEN UP IN A DRAFT!
- Did Alia say it needs to be written up in a draft? (again)

Andy Bierman:
Alia: Can we make rollback not the problem of the network element, to handle managing state per-client?
Andy: Where is the requirement for partitioning by owner?
Ed: Fundamental problem of partitioning the database. Atomicity & state ownership: should be handled separately. Need to know whether we can go with an ACID model. Question is: Can we have atomicity above the RIB manager? Or, do we have to have each of the clients managing that state by reading what is the present state and managing that?

Alia: Question of whether we can punt this to the client or not?
Ed: Caution against getting bound up about how badly people cold hurt themselves with this, because we can do just as much damage with existing Routing Protocols?
Wes George: We don't want this to be a "Nanny Protocol"?
Ram Krishnan: Question about scope? Is scope inclusive of end-to-end of, e.g.: servers in DataCenters?
Ed: Don't know, yet. Ultimately probably depends on what we come up with.
Joel Halpern: Charter is clear. For now, we're focused strictly on routers, (RIB Manager?)
Alia: Servers may participate in some pieces of it, but we haven't been thinking about that. Need to stay focused on the charter.
Ram: Are we considering soft-router or Virtual Switch?

Alia - don't want to rule out use cases accidentally.
Joel Halpern: Going back to rollback, etc. Claimed in framework doc that ops can have start/stop times. If Ops has stop-time, we need to make it go away when state expires. But, other problems. If we say we need multi-headed control, then how can we know what we're doing at any moment in time ... particularly if other clients, at a higher priority, are doing things underneath me at the same time.
Alia: Think this may involve events
Ed: Somebody should look at taxonomy of this space. Concern about state & complexity.
Joel: Same priority ... then, does newest operation win? Rollback either on device or on the controller has a lot of complexity, but its something we need to think about.
Alia - yes it's an important issue. happy for a small group to work on it. good to articulate this space better.
Ed - if we can get consensus for client/controller controller election scheme that'd be great. Better than running some scheme on the back end to figure out who the leader is.

Andy Bierman: Big disconnect between how CLI configures the box and how programmtic interface does it. Right now, we already have situation that 4 diff. admins could login and overwrite changes on each other. In programmatic environment, it just happens much, much faster. You have to do multi-headed control, because one app. isn't going to control the whole world.
Ed - Not necessarily true. May need broker to arbitrate between diff. apps. All problems in CS can be handled by another layer of abstraction.
Andy - Have some experience with NETCONF. Learned from shared candidate that it has these problems.
Ed - Agree with you. Talking about whether it exists on the switch or it exists offboard (somewhere else).
Nitin Bahadur: It should not exist in the switch.
Alia - yes, we're ratholing here.

Sue Hares - I2RS Policy -- "The Basics" Preso
Diagram of I2RS Policy Framework
Policy Framework 101 - Policy definitions / Policy Actions
- Policy: could be implicit or explicit
    - Explicit: what you configure
    - Implicit: What's implied in "protocol" doing the right thing in the configuration

Key Questions: questions related to FIB update. Sue has answered it for 1 use case. For each WG, answer these questions for yourself/your use-case.
I2RS hot-potato client
- read-scope: Interface, FIB; write-scope: FIB

TE I2RS client
- read-scope: same; write-scope: same

hot-potato client selects 1 of 2 exits based on distance

TE client wants to load-balance on same 2 exits
Problem comes with two, multi-headed clients and intersect - could result in congestion outbound on 1 of 2 links
Precedence resolves the issue
Assign different precedence to different clients - one of solutions we could use for this case
Another solution: store if not best. If you lose based on precedence, we'll store it for later. this is about sequencing.
Example: When TE client stops and return to Hot Potato routing
Another example: upon reboot, return to "normal" BGP routing - could revert to BGP routing in config, unless I2RS decides to write something to config.

Key Questions Answered?
- What I2RS clients can talk to I2RS agent? Need identity and security
- What happens if there are 2 I2RS clients? If both try, need precedence. Should precedence be all per-client or overlap?
- What happens if local config overlaps with I2RS state?
    - BGP lost in aboev examples to more specific hot-potato, TE load-balancing, and hot-potato + TE load-balancing
- What happens when I2RS clients remove state?
    - Depends on store-if-not-best flag
- What happens if reboots? Should persist or go away?

Is this multi-headed? Q: If you read on one-side and write on another? Just a question.

Types of commits for I2RS config
- Reserve the TE for this time -or- register start/stop time
- Simple local config gotcha's: policy is both implicit and explicit -- *if you overwrite an implicit policy then you have to put it back!*

How to match this to Andy's talk
- Dynamic configuration -- that's how the I2RS writes the policy for RIB forwarding
- Should we do conditional store?
- Should we have configlets: static, conditional, triggered?

Questions -
Anoop Ghanwani - Multi-headed control shold be kept as a future, because there's just too many problems we could get into.
Ed: Agree. (chair hat off) Could introduce too much complexity and could slow down work if we try to tackle it too early. That said, there are some use cases that may require this so we should talk to.
Anoop: If one guy writing and mahy reading, then that's not multi-headed control
Ed: Agree with nomenclature perspective.
Jan Medved: Multi-headed control -- is that all being pushed up to Controller or not? We need to keep track of state of who pushed some config to NE's. May be needed in the future, even if we don't tackle it now.
Alia: MHC is complicated. Important to talk about it now and understand it, because we don't know how hard/complicated it may be to graft it onto the Architecture later on. Feels like it'll be hard to add on. How much of the complexity do we need now? Are there real use cases that we have now that need it or not?

Or, can we add enough MHC, but don't have to get everything pulled through up front. Think it's good to have those discussions now in Small WG's and try to figure this out.
Jan: Agreed.
Ed - we're fortunate to have some operators here today. would like their opinions.
Himanshu Shah - Don't want single point of failure, so need a solution for that. Perhaps just redundancy of the controllers?
Ed - Maybe it's up to the NE to elect just one controller as "master"
Ed - pretty clear the system has to be redundant.
Himanshu - can be done in a way where both are not actors - so then don't need multi-headed.
Andy: Difference from normal config, e.g.: injecting state to deal with intrusion. Seems like going through one broker when you're under attack, that could be bad, particularly when you have to have that broker remove state.
ed - chair hat off have to disagree. having multiple systems where you have to roll back to previous state is more complicated than single system. Don't conflate logical centralisation with lack of redundancy.
Jan: Difference between multi-headed and redundant. multi-headed problem is different apps trying to inject state and not aware of each other. Multiple Independent Apps is completely different from multiple controllers needed for redundancy. We need latter, we don't necessarily need the first.
Robert Raszuk: want to go back to Ed's single master comment. Single master per app or per network. Per network won't work.
Ed - so you've just said OpenFlow won't work?
Joel - Single, active master per device per I2RS agent, but don't want a single controller for the whole network.
Jan - Implying intermediary that clients connect to?
Joel - Seems sensible answer.
Jan - Agreed.
Alia - glad to see so many volunteers for taxonomy session!
Andy - What matters for MHC is what happens during contention?
Ed - If someone has ownership of that state, then this problem becomes much easier to solve.

Back to Sue Hares preso - "How to match this to Andy's talk"?
Question: Will Simple Policy for other Use Cases Help? If so, Sue will write this up over lunch and send out.

Alia Atlas - Channeling slides from Nic Leymann
- Visions for I2RS as one Building Block in E2E picture
- use case of protecting against DDoS. very selective treatment (only want detected attack traffic diverted to e.g. data center with virtual boxes handling malicious traffic).
- can generalise attack mitigation to "flow-aware traffic steering". So common interface to address multiple problems. A form of service chaining. Traffic shipped to DC and can go through various boxes there.

Ed Crabbe - **Network Modelling Discussion**
- How far down the rabbit hole we go?
    - Have use case on table for network topology discovery
    - Production of SRLG info.
    - Question that hasn't been addressed is how far down to the network component level we go?
        - Physical, Logical, Active, Passive components ... packet, Optical, Physical components, etc.

- Quite a can of worms to describe what some of the physical components: hot-huts, conduits, manholes, etc.

- Simple modelling example on Layer-3: LSP between them. Keep track of interfaces that LSP tunnel lands on.
- Once you add OTN underneath that it gets much, much more complicated. OTN components isn't that hard. However, it's when you get down to the physical component level (inside plant, outside plant) ... it becomes impossible.
- Very simple inside plant example
- Outside plant is even more complicated
- Ed Crabbe: Me, at microphone, being scared of how far down the rabbit hole we want to go.

Adrian Farrell: Personally interested because he wants to see topology database containing this information. However, looking at the charter, it's about "Interface 2 Routing System" and a lot of this information is not in there. For example, media, bandwidth, SRLG, etc. Ultimately, want a core set of information related to routing that we want in there, but don't want us to go too far down the rabbit hole.
Ed - G.709 extensions that are part of PCE, etc. and lines are getting more blurry all the time.
Wes - Q from Jabber (GPONdave) - why do providers want to give detailed information on physical components to others?
Ed - Info. is for use within a provider's network, not for use by external parties.
Joel - Agree with comment to stay away from "Inside Plant". Re: optical ... many others have done a lot of modelling work. If we go there, then we need to make sure that our models are consistent with what they have already done. One other thing is: you forgot "Ethernet". Not optical, but carrier that's below IP/MPLS. Need to concentrate on things that are visible to Layer-3. Don't need to model the whole world, but need to make sure we capture enough to make it useful to Layer-2+.
Himanshu - Agree with Joel. Didn't talk about Layer-2 and we need to model that.
Ed - If we're not incorporating overlay meshes into this work, then we've done something wrong.
Wes - Agree that we don't need to model everything, but need enough so that the "routing system" can make sense of it and make proper decisions (wrt capacity, restoration, etc.).
Ed - Agree. As individual, if we don't have something that can do optical SDN applications, then protocol will become less relevant over time.
Wes - No need to replicate what other protocols already do, but need to be able to re-use them.

Consensus in the room seems to be that we (I2RS) could go down to some part of the optical layer might be OK; however, anything southbound of optical layer (e.g.: physical components of inside/outside plant are out-of-scope).

Jan Medved - Topology Use Cases:

collected these along with bunch of co-authors. Couple of drafts (draft-amante-i2s-topology-use-cases and draft-medved-i2s-topology-requirements). Mix of Cisco, Juniper people plus Shane.

topology use cases introduces a whole bunch of use cases. will go into some in detail here.

framework:
when started looking at this saw need for data model that is presented to applications. Part of work of this WG is to come up with that model. Can go into that in a small group here maybe? Need to figure out required elements for topology data and information model. Comes from multiple information sources - e.g. mining from network elements. Ideally want most of that available via I2RS not via other vendor-proprietary

(but with commonality) protocols / data models from elements (via netconf, SNMP, TL1, CLI etc)

Robert: Assumption - limiting this to single-provider network?
Jan: No.
Ed - Why do we want multi-providers up-front? Going to slow things down, but will need to work on it, eventually.
Jan - Had discussion around single- or multi-provider. Agree to start work on single-provider and step into multiple providers down the road. So, need to make sure this is extensible for multiple provider.
Adrian - (hat back on). charter says “currently limited to single admin domain”. but as Jan says need to be forward looking. so need to be aware of future.
Jan - single domain can be single-AS/multi-AS. I’d prefer to limit to single-AS for now. Start simple.
Sue - +1.
Fabian (NEC) do you consider devices that don’t run I2RS in your topology.
Jan - I think yes. but for this group our work is to come up with NE model that NE has to provide (with I2RS) for topology gathering.
Fabian - reason I ask is this needs to be annotated with info such as “controlled by I2RS”, “not controller by I2RS” etc.
Joel - Pedantic point. At this point, we’re interested in info. model, not data model.
Joel - We don’t have to start with data model.
Jan - Eventually, need to get to data model to know what you’re doing is right.
Alia - We’re chartered for info. models and we’re doing info. models. If you start with data models, you make assumptions that drive back into info. models (that may be wrong). During 3rd group tomorrow is brainstorming topology info. models. Those have interest in working on that, then help out with that effort.
Ed - While developing data model first puts constraints on info model, we can’t stop anyone from doing that. We need to do that in the end. Our goal as a group is info. models.
Jan - Agreed.

Jan - Need partial topo information that NE’s know about, but also need info. from Inventory Collection & Stats Collection system. So, interface is used to get info. from those systems. Those interfaces probably won’t be part of I2RS or solutions developed here and may be based on WebSockets or something else to get that information.
Jan - Also need Topology Server to have interfaces to Policy & Orchestration Manager.
Use Case #1: Capacity Planning
- Data is not normalized and lacks sanity checks on data being input. Inventory data is being updated very infrequently.
- Solution: extract info from NE’s and other sources, e.g.: inventory systems, stats collection. Ultimately, need to get overall (normalized) view of overall topology.

Use Case #2: PCE - a.k.a.: BW orchestrator
- Get info out of Topo. Manager and feeds that into BW Orchestration Manager.
- This is used to answer “Demand Admission Queries” by BW Orchestration Manager.
- BW Orchestration manager is used to drive changes to underlying NE’s using PCEP.
- Use cases requires mostly live info. from network; gets that from BGP-LS. But also, requires statistics information (which isn’t in BGP-LS). So, need to use SNMP for that, because BGP-LS doesn’t carry that.
- BGP-LS is push-based, pushing info. up into Topo. Mgr. Residual BW, Avail. BW, etc. can be carried in BGP-LS and pushed into Topo. Mgr.

Sue Hares - BGP-LS brings fluctuating amount of info from Routing System ... Where does I2RS fit in this? Getting a lot of data up from BGP, via BGP-LS.
Jan - Dampening needs to happen at Topo. Mgr or PCE. Whatever changes occur in network, we do want to see in TED (in real-time).
Sue - Where does I2RS fit? I2RS agent?
Jan - I2RS agent is in NE's. Hope to create common info. model in the NE to export that information northbound to Topo. Mgr. Not sure if I2RS will play between Topo. Mgr and BW Orchestration Mgr.
Sue - Will I2RS Agent be associated with BGP-LS on NE?
Jan - The thing that turns on/off BGP-LS on a NE could be controlled via I2RS.
Jan - BGP-LS doesn't go through I2RS Agent on NE's ... BGP-LS is a routing protocol used to carry dynamic routing information from NE's to Topo. Mgr.
Jan - BGP-LS could be a data stream that is used as input to I2RS.
Robert - you mentioned the TE metric in the PCE use case. Does the PCE also work with segment routing (which doesn't require signalling)
Jan - yes it will.
Nitin: Touched on statistics. Don't think BGP-LS should be used for exporting statistics northbound. Auto-BW is still used by a lot of people. Still could have BW Orchestrator feed off Auto-BW statistics. Haven't talked much about statistics and how they are pushed/pulled out of NE's.
Jan - for statistics we can have topology manager polling. Preferably have I2RS IM as to what stats are needed. Or can have push mechanism where selected set of stats is pushed up to the topology manager when it's important. Need to be careful using push so as not to overwhelm the CPU. tradeoff between push and pull. one mechanism for push it doing selected (small) group of stats and where you push when stats change. BGP-LS does that. that's one possible data-stream used by the topology manager to get a selected group of stats
Nitin: Still don't think BGP-LS is the right thing to push that.
Jan - Ones that are already in IGP can be ...
Fabian - I2RS wants to push things up to Topo. Mgr, right?
Jan - it's both pull and push.
Himanshu - i was thinking not of BGP-LS but of using a mechanism defined in I2RS. And use that for everything - inventory, statistics etc. I agree with Nitin that we should have a common language for I2RS communication for different purposes.
Jan - Need to differentiate between info. models that we need to define and the protocols used to ship that information around. If there is already an existing protocol that meets those needs, then why not use it?
Joel - Some discussions around Topo. talk about abstracting topology. We need to first talk about what info. model needs to look at, then figure out what protocol to use. Similar on events.
Shane - if you wind back to the previous use-case it's important to note that there are 2 different elements in the one draft. (1) TE - short timescales for statistics and changes applied. (2) much longer duration use-case for capacity planning. When talking about Information Models timescales matter. So we may have different models for the 2 use cases. And may use different protocols. it's too early to say. so let's focus on what we need to view from an Information model perspective.
Ed - clearly an allusion to a rich database of information underlying this thing. Lots of this only of use to NE in abstract manner (via SRLG). NE's won't maintain underlying topology unless we've made a mistake.
Andy - problem is one of harmonizing data and info. models don't help there.
Sriganesh - is there one toplpogy manager with different push and pull models or are there multiple. and if the latter how do we reconcile them.
Jan - in framework have one that can reconcile information from different sources
Sriganesh - scaling issue?

Jan - can have multiple instances talking to each other in different domains.
Sriganesh - seems a stretch to have one topology manager even for one domain.
Jan - not really. can scale using clusters etc.
Chris L. - you took the words out of my mouth. it's a centralised function that can be logically distributed or clustered.
Jan - lots of us build this stuff.

Jan - ALTO Server - Another Use Case. Topo. info. is delivered ito ALTO server. ALTO server feeds multiple clients. ALTO, in my mind, is a modelling language. ALTO is one way to express topology in abstract way. It's been around a while and, perhaps, a good one to use.
Joel - Is Topo. Mgr API really the I2RS. Is how the Topo. Mgr gets its information is out-of-scope and I2RS is really focused on the upper-layer API from Topo. Mgr northbound?
Jan - Yes.
Himanshu - Draft talks about I2RS agent in NE.
Jan - Actually, we're talking about I2RS above Topo. Mgr.
Himanshu - Doesn't I2RS reside in NE.
Jan - Yes, it does.
Himanshu - is the topology manager an I2RS client, or is the ALTO server the client.
Alia - NE does not imply a router. It's possible for a I2RS client to provide information from down below and an I2RS Agent to pull information to above.
Chris L. - Example could be a VPN configuration point. I2RS talking down to NE's to realize a configuration, but abstracting the topology northbound to apps that need logical view.

Use Case #4 - Service Router
Use Case #5 - VPN Services Provisioning
- Instantiate new services at the appropriate places, validate ACL's are configured properly. Topo. Mgr finds PE's with appropriate uplink BW, access circuit type/media, access circuit capacity, etc.
Use Case #6 - Troubleshooting & Monitoring

Ed - Introductory slides are uploaded to IETF Web Site. Break for lunch.

4:08pm - Reviews after Small Group Breakout #1

Multi-Headed Control (MHC) WG Review - Sue Hares
Question: Where is the complexity?
1. Multiple clients per broker?
2. Apps coordinating with single edge? NFV, service-chaining
3. More knowledge at the edge
4. Combination

Multiple clients per broker?
Apps Resolve the Issue or Precedence Resolves the Issue?
What happens when you have 2,000 clients?
I2RS Datastore Framework
Conditional Configuration

Jan - With broker model do we propagate information from ultimately source down to the I2RS agent. Or, do we just keep state in the broker?
Sue - We propagate state down to the I2RS agent.
Dave Hood (Jabber) - How does Agent #11 know the commands are in conflict or are intentional?
Sue - Agent 11 is just going to note that the two clients have told it to do the same thing.
Sue - You have to tell a client what you're going to watch.
Andy - One of the big issues is: what is a collision? How does the server know that these 2 clients are causing the conflict? At the lowest level, the server will know that two clients have tried to access the same data. But, then there's hierarchy of the data at the server and dealing with that.
Ed - Keep going back to precedence. In a well-behaved network, the only reason that MHC should occur will be when you have something go wrong and 2 apps are not cooperating. Haven't seen an actual use case where this will occur normally.
Sri - Diff. clients with diff. capabilities. Clients has least amount of state; however, broker will have full state so it can make change more intelligently.
Dave Hood - Suggested principle. Joint ownership of resource is not allowed. Switch just does what it is told. Whether the switch has ownership or broker is to be discussed.
Sue - If you have no joint ownership, then you have no collisions. Fine for the same protocol, where this shouldn't occur.
Ed - Multiple protocols at different priorities could be where collisions occur.
Adrian - The broker is a SPOF. What if I run 2 brokers for redundancy? How closely do their collaborate? Or, more specifically, what happens when the two brokers can't communicate to each other.
Joel - This is a solved problem. Other folks have figured out how to solve this problem.
Adrian - Why not assume there is always 1 client to 1 broker?
Joel - Clients are assumed to not be coordinated. Broker is assumed to be coordinated. Unreasonable to expect clients to all be coordinated. Even in routing protocols we don't deal with how do we get routing protocols to synchronize properly.
Adrian - Given the range of clients. Could you expand on that?
Joel - Given the set of use cases, there could be a large set of machinery that could rely on this architecture. We can't assume that the clients will be properly coordinated amongst each other.
Dave Hood - One or more clients may not be expected to be reasonably coordinated, because a client may be owned by another administrator.
Ed - Need to have the same protocol speak to the same controller. Don't know that we need to have protocol that speaks East-West to other controllers in a standards-based way.
Andy - Do we have consensus that the wide range of clients that resources could be taken away from any given client at any given time. Think about scheduled maintenance.
Ed - Broker could easily take away that state from the NE's, because it knows that it doesn't have state that any client(s) are bound to it.
Joel - Things can happen: links fail, boards crash, etc. We're going to have to say from broker to client that "we know you asked for this, but it's not there any more". If we can move the arbitration out to the broker, we make our lives a lot easier.
Sue - What about the CLI?

BGP WG Review - Robert Raszuk
Use Case #1 - Troubleshooting & Analysis of BGP - tried to find out what things couldn't be done with tools today. Things like: route reachability, expected exit point, route hijack detection, route stability, dropped routes (policy, malformed, etc.); reporting dampening/instable routes; protocol statistics
Applicability: NOC troubleshooting, reactive & proactive reporting, increased network stability

I2RS BGP Analyzer:
- BGP Analzyer Mode: 3 options
1. On-demand querying for a prefix -- query each node in the network on-demand
2. pub/sub model -- only listen for events
3. database based model -- record all BGP UPDATE's
- Communication Channel: MRT, BMP, Programmatic API's
BGP Analyzer Scale:
- Stats and prefix events require analyzer to scale to millionso of prefixes and paths
- Scale directly proportional to current prefix churn & BGP routers in a given network = ~100 prefixes/sec

Ed - Non-Loc-RIB routes you have some BMP-like functionality and recording it all in some offline DB -or- your recording this in all network elements. How is this not like BMP with filters?
Robert - Yes. That would be OK.

Use Case #2 - Performance Based Routing
- Compute least delay exit paths, least-cost exist paths, assure SLA's, reduce jitter and RTT of dataplane, spread utilization of external links, enhance BGP TE
- Applicable to ISP networks & customer networks
- BGP Monitor:
    - Pub/sub model for routes
    - Dataplane statistics
- Communication Channel: Programmatic API's
Shane - Dataplane statistics: what does that mean?
Robert - Could be interface counters on routers
Shane - That's important, because it could be statistics from the NE's themselves or it could come from Statistics collection warehouse.
Nabil - Where does an I2RS fit in?
Robert - Collection of information or stats from NE's. Distribution of decision to NE's. Can argue for NETCONF, or to I2RS.
Nabil - so API is to influence BGP, not to do the analysis? Need to call that out.
Ed - Nitin/Wes have use case where you can do that by altering the RIB.
Robert - if you have 500 routers in the domain it's much easier to configure one (and distribute policy with BGP) than to configure all 500.
Nitin - where does I2RS come into the picture? If you're modifying BGP then why do you need the I2RS channel?
Robert - it's an option. You could use NETCONF.
Shane - to answer Nitin. We need to think about the fact that there's no standardised interface into BGP today.
Ed - so what we want is a Yang model for that.
Alia - NETMOD exists for a reason. We're doing the cases that are needed for I2RS. Others can be done in NETMOD.
Shane - back to Ed's point you may still want to have I2RS controller be able to make multiple simultaneous writes to different NEs even if you have BGP pushing stuff around. I don't want more interfaces into my NEs. Want to leverage the ones I have more and more.
Alia - I want to focus on the interfaces that need I2RS functionality rather than looking at where we help NETMOD.
Nitin - we could look at Nick Tingle's presentation on "what is wrong with BGP" (re IXes).
Robert - that's not in scope of I2RS.

Adrian - puzzled by what Alia just said re the difference between “things for I2RS” and “things that require modifications to NETCONF data models”. Here we’re talking about what we need for I2RS, not about how to achieve it.
Alia - if there’s a use case to do this with BGP for NETCONF then don’t see why it needs to be in I2RS. I’m not trying to rule out use-cases that don’t require stuff beyond NETCONF but that isn’t our focus.
Adrian - you need to rule out all the cases that can be solved by what we have already. Focus on where we have pieces missing. People seem to be saying “I2RS is a protocol”. It’s not. We want to leverage what we already have. But clearly we’re missing pieces (e.g. the data model for this stuff). Those missing pieces are I2RS.
Alia - we can take NETCONF and warp it to do what we want. or use FORCES. but there’s NETMOD. NETMOD is trying to do data models for NETCONF. If they’re not doing BGP models for lack of input then...
Adrian - if you’re saying “where should we write the data model” then I don’t care. If you’re saying “where do we work out what data model we need to write” then that’s here.
Shane - don’t think we should focus just on what can’t we do that we need I2RS for. Also need to look at how to make current network operations more efficient.
Alia - I hear you. I’m trying to avoid boiling the ocean. Just want to get some good solid use-cases done.
Thomas Narten (jabber) - I2RS should focus on areas where there are big gaps. if 80% of work for a use-case has been done elsewhere then why bring it here? finish the work elsewhere. So pick problems that are big enough that I2RS is needed to solve them.
Robert - Some of these things may feed to GROW, as well.
BGP Perf. Monitor Scale (see above)
Use Case #3 - Don’t do this in I2RS and use NETCONF instead
- Centralize BGP Policy
- VPN Provisioning
- VPN Stitching

Undecided -- not clear if this belongs in the protocol or outside of it
- BGP Error Handling
- BGP Route Manipulation
- BGP Optimized Exit

Inline in BGP
- BGP flowspec
- BGP RT filter -- already exists in BGP, but may be needed for legacy equipment that doesn’t support it
- BGP Optimized Exit
- BGP Operational Messages -- currently, in BGP ... so, should still be in protocol

Use Case #4 - Service Chaining
- Use of next BGP attribute: Next-Hop chaining
- Each NH corresponds to a different next-hop in the chain
    - Control plan based solution compared to data plane scaling
    - More scalable compared to existing source-routing approaches
- Applies to any AFI/SAFI
- Works Inter-AS/Inter Provider

Nabil - Explain a little bit of next-hop. Is it ordered from the service points.
Robert - The attribute is ordered.

Nabil - It's all encoded in one attribute. And, all the attributes are the next-hop in the chain.
Robert - It's like a loose ERO.
Keyur - Attribute will also cover any tunnel encapsulation.
Nabil - Could also leverage the underlying tunneling mechanism of the VPN. How complex is that to manage? It's simplified it a little bit, because it's the service point that's encoded. The next-hop in the chain is encoded in the attribute.
Robert - All info. is distributed everywhere in BGP already ...
Nitin - Does this assume that each node in the service-chain needs to speak BGP?
Robert - No. This assumes that NE can reach the destination node. Service node could just have a default route.
Nitin - If you want to do service-chaining in my DataCenter, then I need to run BGP in the DC. Need to run BGP on my internal nodes.
Robert - Not if I'm using tunneling. Will take it offline.
Luyuan - This is a BGP enhancement, not exactly I2RS.
Robert - Programming service hops can be done centrally, so it's here. If service node is not running BGP, then you can use I2RS to program the nodes that don't speak BGP.
Nabil - Need to determine scope of where this applies. Are you intended to encode the service in the attribute.
Robert - Encode the svc ID in the attribute or just encapsulate that in the next-hop. Just generate next-hops per service.
Nabil - Need to know the context of a packet as you're receiving it. Not sure how this would work in a virtualized environment. Why is this not just policy-based routing?
Robert - You could.
Keyur - Either you install routes if you don't run BGP. You need routes there -- I2RS does that or BGP distributes that for you.

RIB Use Case - Wes George / Nitin Bahadur
Fair amount of this space that is already solved. Pieces that allow you to talk to the RIB today. We looked at how to make this cohesive & comprehensive.
Different RIB's per different IGP that are distilled down to one RIB.
Things you need form the RIB:
- routes installed
- candidate routes
- RIB Tables
- Read-only data from *FIB* -- perhaps you care about available of forwarding entry, i.e.: if it was successfully programmed in FIB or not. Additional level of verification

Alia - Talked about events when we talked about this before. Treat changes to FIB entries as events, rather than having to read them all the time.
Nitin - If the router reboots, then you need to read it from the FIB. Going to RIB Manager and ask it for what entries are installed in the FIB.
Alia - Concern was if we have to go down to LC's to figure out what's in FIB. Fine with going to RIB Manaer to just ask it what's in the FIB.

Assumptions
1. Controller is broker between apps and network -- don't care if controller is distributed or centralized, at this point.
    - Deals with prioritization/conflict

- Signals pass/fail to app requests
- Provide info to apps

2. Controller has different RIBs for each AF
3. Controller has only one RIB per AF
4. Clients have different RIBs for each AF -- while Mutli-Topology does exist, don't want to deal with it for now.
5. Clients have different RIBs within AF -- different IGP's

Jan - Per AF, per NE?
Wes - Yes.
Jan - Is information model in NE the same as informational model as in the controller?
Wes - We didn't talk about that.
Nitin - What's exposed controller to apps: don't have a firm idea on if it should be the same or if it should be different.
Jan - Points for and points against. If you haven't thought it, yet, then probably should note that and we'll talk about it later.

Diagram of I2RS agent
- map/select traffic to install entry in RIB and communicate from RIB <-> FIB.
- install in RIB set of next-hops; set of NH's may be IP; may be MPLS, could be backup NH's

Building Blocks
RIB potentially has interactions with optical layer. Not proposing I2RS to control optical layer, but may not be limited to IP network (e.g. if optimising overlay - underlying transport layer may have an optical express path you can use). recommending an interface to the system that controls the optical network where you can articulate your requirements (latency, bandwidth, time-bounds, SRLGs) and optical network can tell you if it sets up a path. Increasing overlap between IP and Optical now we have optical layers that understand Ethernet etc. This is an extension. sometimes optical network ends up being part of the routing path.

Nitin - This is probably also a function of the PCE. If PCE wants to optimize function in the network and wants to talk to optical control in the network, then it can do that.

Optimizing Exit Control
- Traffic could be from server or from a router, want to control which exit point it uses to egress the network
- Various ways to do this
- Could use BGP, manipulate those atttibutes
- Based on business attributes it can program multiple things all the way down to the transport level. Or, it can say that it just wants to tell it the next-hop or the edge interface to leave the network. Abstraction could be MPLS label, GRE header -- and at egress point, it recognizes the incoming header/overlay and strips it off before transitmitting it on to ultimate destination.

Q&A:

Fabian - this looks awfully close to a non-Ethernet OpenFlow. Is this still routing?
Nitin - good question. This is still routing, not forwarding. You can say "I want traffic from this server to go out of that edge router". You can look at LSP or IP path to that egress. Not programming a match action.
Wes - not to say you couldn't use OpenFlow for some of the same things.
Ed - I agree, but it's a very blurry line.

Luyuan - what exactly are you proposing that's new?
Wes - sure, a lot of these things can be done today (gives examples). I2RS gives clear picture of the whole thing such that you have better information to make decisions.
Luyuan - more of a case of formalising it?
Wes - yes, being able to standardise interfaces so you have common control point.
Nitin - more than that. With PCE can't specify end point. Also require PCE-P which servers may not have.
Ram - we covered this in our group (service chaining).
Keyur - this isn't service chaining, it's building an overlay
Nabil - a service chain is an overlay.
Wes - service chaining is part of discussion but it isn't all of it. Need to co-ordinate.
Alia - this is RIB manipulation. Service chaining is a use-case of how we can use RIB manipulation.
Nabil - yes, many use cases have shared mechanisms. Some cases are also degenerate cases of others. need to separate APIs from all the stuff that goes on top. Need to avoid rehashing SDN overall and focus on the OpenFlow equivalence in talking to RIB or control plane. Otherwise this is SDN deja-vu all over again.
Alia - we have lots of rehashing. we need to refocus. One way to get details is that in next small group we boil this down into what data you are reading/writing so we can get to the information models. I'm concerned we're trying to limit/define what goes above I2RS instead of focusing on the mechanisms and tools it provides.
Nabil - better to step back and say we're fitting in an SDN paradigm. These are the interfaces we're defining. Then focus on the interfaces.
Jan - have you considered adding ACL programming to the RIB model? Might be useful.
Nitin - let's discuss off-line.
Wes - mostly when you use ACLs to filter you're dealing in terms of setting up dump paths in the network. Would make sense to have that as a part of this. So you can say e.g. "I want to put this ACL at all my edges". So it's a subcase of DoS mitigation.
Jan - did you consider value add tools like checking for forwarding loops. Controller can do the checking that the information the apps want to prorgrm is valid
Wes - sure, but up to user as to whether to have controller do that or leave it to the apps.
Nitin - don't see this as something IETF needs to standardize.

Ed - we're going to punt topology to tomorrow. Start 9am tomorrow with donuts.

4/23 - 9:10am Start of Second Day

Topology Model Discussion - Joel Halpern
Tagging for each layer, in order that (for example) a PCE knows what layer of network to perform a path computation on.

Nabil - Is there a relationship between different (OSI) layers?
Joel - Relationships extend up and down:
- provides services to (e.g.: server layer)
- dependent on service from

Nabil - Is relationship between layers in scope?
Joel - Yes.
Ed - Same question and I think that it is in scope.
Joel - This is vitally important point
Ed - Need to figure out how to model a port that provides service to multiple services/layers

Ed - Multiple logical instantiations of a single physical port.
Joel - Yes, that's probably the best way to do it.
Ed - I don't know that this is the best way to do this.
Nitin - donuts are good :)
Ed - If you have segments that contain multiples, then you need an ordered list.
Joel - Ordered can come from recursing down the layers of the graph. There was debate in the Small WG as to how you represent the ordering.
Jan - Do you replicate the info. in the DB or does your query actually ask for (do SQL JOIN) to establish what is the order/dependency.
Ed - Making a convenient model is important.
Nitin - Considered multiple services at the same layer, e.g.: network-layer that could be a terminating point for VPN service and multicast service. Could be different services, but terminate on the same node.
Joel - Yes. Didn't discuss being a terminating vs. transiting point. Yes, things can support services.
Wes George - In PCE/CCAMP, is there already a topology model?
Adrian - How many wheels do you want to reinvent? The model is implicit. There are arch. docs about how things fit together across layers and consequences of things about that in the protocol documents. There is a modelling worldview in Geneva, which is very Transport oriented, but we would be irrational not to read it.
Joel - We want this model to be compatible with that so we could incorporate that model into here and not reinvent it or have completely different model of our own.
Adrian - They just have links or nodes. We should only use it if we think it is "right".
Ed - If it is weird, then we may not want to use it. We should look at it, before passing judgement on it.
Jan - Arch. docs for PCE assumes that its using a TED to compute paths. BGP-LS will transparently pass info from IGP's toward TED.
Andy - Question on Arch. of standard. Heard phrase that all data is centralized? Are we standardizing the controller to centralize all the information?
Joel - That is the question I said we didn't get to. We need to standardize all the information that I2RS is going to receive or is I2RS going to be responsible for transporting the information to a I2RS controller.
Sue Hares - Challenge for Jan: Can we see the information model between all the bits & bytes that BGP-LS are transporting toward a I2RS Controller.
Jan - Excellent point and I take that as an action item.
Ed - If people are going to want to use this model for design-rule checking, then it needs to be amenable to that.
Joel - That was something we discussed, but didn't make it into these slides. Primary point of slides was the abstract information model.

Question that Adrian brought up: is this model only used to get information *from the topology* -or- does this model assist with making changes to the network?

Alia - Charter says we can only create read-only information model, we cannot make changes to the topology.

Seems that there may be disagreement between AD's, co-chairs & charter on whether the information model is read-only or read-write in terms of the NE's/topology ...

Nabil - Passive elements, you mean inventory that is not turned on.
Joel - No, could be stuff that you cannot discover, because they are not active, (patch panels).. Could be that, though.
Nabil - There may be value in discovering things like SFP that aren't powered on ... but, we need to tighten

up the definition of what's active and what's passive?
Joel - Yes, need to do that.
Wes - Talking about I2RS, so we should draw the line about excluding things from lower layers in the network. We should expect that them to automatically export that to higher layers.
Joel - Can't do that. Even an Ethernet switch could be considered "passive" to this model, if it's just providing Layer-2 transport.
Wes - We need to, perhaps, represent something like SRLG, but not more details than that from layers below.
Joel - Need to represent both horizontal and vertical dependences between different layers.
Nitin - Be careful to not convert this into Inventory Management system.
Shane - Use cases incorporate topology info for things like L2 switches in DC's or Switched MetroE environments, etc.

Ed - there are a few characteristics of a modelling language that would enable it to be extensible to an arbitrary degree. doesn't mean we need to define it all up front.
Andy - maintaining this kind of database is expensive. having one per application may not fly. Of course we're not modelling fans/PSUs but what if you get fan/PSU alarms. So it's hard to say "I never care about this thing because it never affects that thing".
Joel - we can extend the structure to express the whole BoM for a router, but I don't recommend it. that's not the problem space we're chartered to solve.
Jan - need to define a model is extensible, but have to start with something that is simple.
Joel - Yep, that's the page we're on.
Jan - Fans, Power, etc. -- that may be useful when you want to optimize power efficiency on the network.

Other Topology Model Use Cases? Service Chaining, etc. -- Giles Heron
Mechanism to learn and to provision:
1. Customer topology
2. Service topology
3. Service chaining
... starting from AC, Ingress PE, interior switches that have tunnels through to Egress PE
I2RS API would touch on Ingress PE, interior switches, etc.

Joel - Are we assuming service chaining is using tunnels? I like that approach, but some may not.
Giles - Tunnel may be a VLAN, not just a MPLS LSP.
Joel - Some have proposed reclassifying the packet at each hop.
Nabil - Have later slides that talk about that later. There are two approaches. Either classify at ingress or at each hop.

Functions we need:
1. Testing for liveliness
2. Performance - monitoring and stats
Service Instantiation
Measure node capacity: CPU, BW, etc.

Service Chaining
Model 1: create whole chain at ingress -- BGP WG talked about that yesterday
Model 2: recursive model where you re-classify at each hop -- problem with this is you have to provision each node in the chain a priori.

Need to support Mux/Demux
Mux example - single FW, multiple users
Demux
Scope & Scale: applicable to SP edge and NFV in DataCenter.
3 Types of Operations:
1. Discovery/publishing when resources changes.
2. Configuration - when setting up a service.
3. Monitoring/statistics - very high frequency (should these be passed over I2RS API?)

Nabil - Could be capability discovery as well. Node may announce what are its capabilities: what services it can provide, but also what are its capacities as well.
Giles - It could be a PE announcing what are its LC's and whether their generic service blade.s.

Alia - Talked about ability to configure OAM, but not clear on where that ought to go. We're talking about multiple channels across a transport, e.g.: statistics, monitoring, etc. that are set-up for different reasons. Whether I2RS are set-up as a single protocol or multiple protocols is still unclear.

What is different for I2RS?
Typically this is done via CLI or AAA systems, e.g.: RADIUS
Standardization of data/information model
Need to get rid of CLI by standardizing those models to make them more machine-friendly.

Linda - Listed multiple topologies, can't they be 1 topology?
Giles - Yes, they could be the same thing in our example slide.
Nabil - There could be null states in the service-chaining.
Linda - True of network service. Could be more complex.
Alia - The other 2 topics that got pulled into this one were customer ports and how they got pulled into the topology model. One of the questions we've heard is: why can't an existing protocol reveal them? Need to know who those customer ports are connected to, etc.
Nabil - Meant by customer topology was a virtual topology for a VPN.
Alia - Do you see service topology and customer topology as different?
Nabil - Yes. Could construct multiple service-chains for a single customer. Service nodes could be different from the network nodes.
Alia - PoP's service topology and virtual (network) overlay topology.
Alia - Customer ports showing up in the topology and peering ports in topology model and thinking about why we want information in there.
Giles - Yeah, we didn't address that. Assumption up to now is that all port info is in the IGP, but every sane SP doesn't put those ports into their IGP.
Jan - Multiple topology models: one is the base IGP used for TE, another one is BGP and another one is knowing where the services are located. The latter is where we move away from pure network topology to service topology.
Linda - ???
Fabian - Topology is missing crucial point, because you need to classify what service policy or service chain that it goes into.
Giles - Yes, that is the forwarding context, e.g.: on the Ingress PE, and then the packet is colored with some metadata carried through the network
Fabian - Forwarrding context needs more information than just routing information. Forwarding context needs to be much richer than just routing context.

Giles - Yeah, needs to be more than just L4+ state stored in packet headers.
Ed - Time-based? Do you need that to be in the routers/forwarding nodes themselves? Or, can that all be in the controller?
Joel - What I don't like is that the I2RS agent needed to maintain connectivity to all controllers and if I lost the connection, then I start tearing down state in the network. That is bad.
Ed - But, you have things like RR's.
Joel - But, that's considered a serious failure in the network. If you're unwinding state in the network, because you lost connectivity to the controller is wrong.
Ed - I'm not saying that we should not have soft-state. What I'm calling into question is when you have diurnal patterns that are truly time-based, then do we really have use cases for that?
Himanshu - Reason for provision and de-provision is to dynamically instantiate a video conference.

Ken (Gray?) - It's good to be able to specify temporal for FW policies that need to be in place during workday hours and for that to be in place on the network device. That's helpful if you go to that device directly and are looking at what is its state.
Ed - That's a philosophical debate.
ChrisL - Making an assumption that that FW is always handling that traffic. What if another FW is handling that traffic? In that case, I don't want those rules in a device that isn't doing something with it. Better to push temporal state.
Alia - Not looking at a pure controller-only model. Starting with a distributed control plane and working on a hybrid model where the controller needs to cooperate with the controller.
Ed - Not germane to this discussion. This is more about
Alia - State duration is important. Some people may not want to tie state to controller liveliness, others may want to teardown state when you lose connectivity to controller.
Ed - If folks have requirements or use-case for state/service installation based on ToD, then they should speak up. Separate discussion on temporal state installed from a controller.
Jeff Tanstura - Could be very "expensive" to install new state, particularly when you're talking about installing state down into ASIC's.
Joel - Liveness triggering related thoughts about Multi-Headed Control from yesterday. If we have multiple applications with always-on controllers creates this problem.
Ed - We've entered into a philosophical discussion. Not all features are going to be designed into I2RS. Some have in mind: always-on controllers, always-off controllers; priorities and precedence; multi-headed controllers or not; etc.
Alia - Need to connect use-cases to architecture, in terms of what is needed.
Giles - Need to tease out the options that are in our use cases and what are the implications.

2nd set of small-group presentations

Wes - AAA

have notion of simplified model with one admin domain with one group. then multiple groups in one domain (so less trust etc.)

Jan - ? missed question ?

Wes - not clear to me if we can take any existing AAA protocols and support this model today. RADIUS vs TACACS. Venn diagram - both good at some things. Overlap of stuff they're both good at. Need a AAA expert to look at this once identified requirements. Extend existing protocols or define a new one?

Bi-directional. App must trust controller/NE. Controller/NE must trust app. Probably need key/cert-based system with all the usual mechanisms. nothing new there.

Set of standard roles. similar to AAA for network interaction but possibly more granular (wider range of applications than of users). Trust implies roles/capabilities etc. Need hierarchy/inheritance. So if you want to let an app access BGP might inherit from a general routing protocol access which might inherit from general config access etc.

Andy Bierman - I did a lot of work on NETCONF access control. Defined in terms of what you can do to what data. Would we need anything beyond data-based access control? Or is it too early to tell?

Wes - not sure what you mean?

Andy - data is named. If data is under BGP then you can say which users/groups are allowed to access that data. Does I2RS generate any new requirements or will the usual stuff work?

Wes - answer is I don't know.

Wes - breaking up into pieces. Read first. Define what info you can get, at what granularity and frequency. Identify push/pull or pub/sub. Define what events you want to be notified about, what triggers you care about.

Sue Hares - what about timeliness? Did you consider use cases of "I need a response within X timeframe"?

Wes - no.

Wes - dealing with identifying areas of config that a given app or class of apps needs access to. if an app interacts with BGP it probably doesn't need triggers on interface counters but might want interface up/down triggers. Also may want to limit how far the app is allowed to go (so if a knob goes to 11 might only let the app turn it to 8). want a set of safety valves to control when state gets out of hand. Have precedence/locking issues. Don't (generally) want app to lock the entire config when making a change. So want to be able to identify a subset of the config (e.g. when modifying routing config you don't need to lock the interface config). could either be the app or the controller identifying sections that need to be locked.

Ed - this feels like a config paradigm. When you say lock do you mean "lock the config" or "lock the RIB"?

Wes - that's why we've talked about single control. Might only want single access to the RIB.

Ed - not sure how things like rollback would interact with the RIB specifically

Wes - when directly programming the RIB you'd want the ability to roll back those changes. when multiple apps talk to an NE or controller you need to manage the rollbacks/diffs based on security rules you've defined. it's a variant of the multi-headed problem. do you have authority to roll back all changes, or only your own.

Ed - we really don't want to have to deal with complex situations like that.

Wes - not sure system will have ability to deal with that. but it may have to flag what has happened.

Ed - we need to articulate this so we can avoid it.

Nitin - this feels more like a function of the controller than something we need to expose over the I2RS channel.

Ed - I agree with you mechanically, but we need to take it into account.

Wes - not making distinction between app talking to a controller and app/controller talking to an NE. Might be you need a different security model. I'd like to see a single model. Need to define it. This needs to be generic so some of it may not make sense in all use cases.

Joel - this may be simpler once we've figured out what we're going to do about multi-headed controller. If there's an entity on top that knows which user did what then it can control that. so you separate out the overlap problem from the pure "roll back to an operational state" problem.

Wes - related issue. Chris pointed out that you may have multiple things using the same application to talk ot the same infrastructure. Do you want pass through authentication so you know which entity is doing what? that extra layer adds potential complexity.

Chris - was thinking of an operator-driven application where the application needs to be authenticated to talk to the broker but also want to grant permissions to the person sitting behind the application. What the app is allowed to do depends on who's driving it.

Wes - other items not connected to reads/writes may be clearing counters/sessions/state/etc. Control who can do that, when and how often. Also stuff related to troubleshooting (bouncing sessions, interfaces etc.) - not config related but needs AAA. Also capacity/scale issues. How many sessions are too many? How much RIB manipulation is too much? How much state is too much? How do you decide where that state is kept? Hard thresholds vs soft ones? (hard = prevent access, soft = give warning but allow).

Wes - accounting too. define who did what when and why. But also log failures (why did they happen). And log precedence issues - what tie-breaker decided who could do what?

Wes - big question of where do you manage state. whether it's app->NE or app->controller you have different types of state you need to deal with. do you need checkpoints? or is it asynchronous and you need updates periodically? Different apps will have different requirements.

Wes - security. Divergence between transport-layer and app-layer security. Do you encrypt payload? Do you do session security? how much do you validate data at app layer or do you trust the transport layer? Reliability - UDP ("I'm going to blast data at you") vs TCP? App acks vs transport apps?

Joel - assuming we want to support a broker model. if we authenticate broker but the user affects what happens the have we got an Oauth problem?

Topology Manager Use Case - Joel Halpern

Target: to be useful, this model has to be exposed via a protocol; but, to be realizable we cannot model everything (down to the physical level) in a model
Specific Use Cases To Solve - Provide topology abstraction that can be used for network-oriented decisions. The use of this information may be actuated by other mechanisms (e.g.: PCE, Openflow, etc.)
1. Failure impact analysis (troubleshooting tool): need to do horizontal and vertical propagation of failures
2. VPN Service Provisioning - needs to know what the underlying network topology looks like, it's capacity
3. Capacity Planning & Traffic Engineering
- Demands may be expressed in terms of observed pairs (city pair) or reservation requests

Graph Element Properties
- Relationship to other elements
    - All are one to many unidirectional
    - Connnect to
    - Used by
    - Uses
    - Service tags
    - Need to indicate direction of failure propogation -- when A fails, that causes B to fails, but not vice-versa, etc. Containment a sub-case of this.
    - Minimum requirements (like # of component-links in a LAG)
- Need extensible kinds of elements
- Extensible properties of defined elements

Ed - Talking about defining nouns & verbs. Can we express design verification requirements in these models? When you talk about minimum # of links in a LAG, that is a (kind of) design verification requirement.
Joel - Yes, that needs more work.

Kinds of graph elements
- Network links
- Network nodes: routers, switches, RR's, service delivery node, ***NOT TALKING ABOUT CHASSIS***! Not going into everything that the entity MIB defines.
- Network ports: Physical, Logical, Nested, Tunnel
- Do we need containers? Probably needed.
- Do we need customers? (Because we're talking about services). Probably needed.

General Graph Element properties
- Identities: need sufficient clarity to correlate to other models, e.g.: SNMP ifIndex to bind a model to a physical port. Corrollary is we may need multiple identities as we evolve or incorporate other models.
- Network location? Geography, POP, IX, etc.
    - Can't standardize nomenclature used for POP, etc.; but, need a container to represent that.
- Administrative & Operational State
- Shared Fate Tagging: separate from failure propogation

Shane - I like the simplicity of what you're suggesting but the reality of network operations might suggest otherwise. e.g. a port may be "reserved" ready for a tail circuit to show up. So "admin up/down" may not be enough states.

Ed - yes needs to be generic as we can't define what each operator's administrative state machine will look

like.

Joel - what you find is you want to be able to say “is this working or not”. if it’s administratively out of service that’s very different to “it has failed”. But don’t want the system to have to know why it’s administratively out of service. you need abstractions to go with it. need to watch out for which use cases we’re solving.

I2RS Topology Use Case - Giles Heron

Use-Case In Context
- Network application here is assumed to be customer service mgmt
- Going to drill down on the L3VPN use case
L3VPN Service Provisioning with I2RS
L3VPN Service Chaining with I2RS - going through intermediate devices in the center of the network
Assumption:
1. SP already supports L3VPN
2. Hybrid model: I2RS + existing L3VPN serivnce
Events needs:
PE/CE interfaces, routing protocols going up/down. Control plane thresholds, data-plane thresholds.
Client pull from Agent
- Some stuff provisioned with I2RS and some stuff provisioned with traditional/legacy mechanisms.
- Clent is going to pull candidate IP interfaces
- What shapers/policers/ACLs are on those interfaces
- Control plane resource availability
- Traffic Stats -- might be pulled from Agent or might be pushed from Agent to client
Data - Client push to Agent
- VRF, RD/RT, interface mapping to VRF, interface OAM, QoS (policers/shapers), ACLs, PE-CE routing protocols
Data - Published by Agent -- stuff that would otherwise get published/pulled via SNMP
Thinking that this is a little bits of information at a time. PE-CE interface might be a lot of info at once.
Client-Agent interaction - Client pushes data down to agent on a per-agent basis; no reqm’t for transactions across multiple NE’s
Hybrid Model: segregate resources what’s set-up via I2RS vs. legacy mechanisms.
NE publishes events

Jan - How is this different from NMS application? Is there more than 1 application/client? Are we replicating existing functionality?
Giles - Comes back to MHC discussion. There may be multiple apps that sit above I2RS client.
Jan - Why doesn’t current EMS/NMS do this already?
Luyuan - Today we have provisioning system and it pushes config down to NE’s. Problem is that each vendor’s CLI is different. I2RS could standardize the API used to push config down to box.
Giles - Other problem is checking resource availability.
Jan - This use case may just be a more flexible/adaptable NMS/EMS.
Luyuan - We’re taking L3VPN as just one example, but this applies to other technologies.
Himanshu - I2RS should not be about just providing new services, its about enhancing ability to make existing more services more flexibility
Giles - Still have question about the applicability to I2RS, since this is about services ...

Jan - Are we enabled new architecture of NMS/EMS? Rather than having a monolothic EMS/NMS from 1 vendor, this could work better for multiple vendors here.
ChrisL - Is I2RS something that allows us to interact with the routing environment or is it a new NMS/EMS model. WRT the latter, we've attempted to do this before in NETCONF/NETMOD. If we are looking at the latter, what are we hoping to do better?
Giles - That's exactly my scope concern.
ChrisL - Hard-state vs. ephemeral state. The latter is something that goes in the RIB; the former is something written in a config. If we're going to do something that's already covered in another WG are we trying to introduce protocol #16?
Alia - Would like to see more use cases where this is done dynamically.
Giles - A lot of the customer-facing stuff might be applicable to those interfaces, peer-facing, etc. interface. IOW, when you generate SNMP traps it should be the same thing. Maybe that's what we should be focused on rather than replacing NETCONF.
Nitin - See us only making forward progress if the service is providing more value. If you're not creating a new VPN service every minute, then why would an SP adapt their existing service to this model.
Giles - We modeled a service we already knew so we had something to talk about, not that we believe that this is what SP's are going to do.
Jan - From that perspective, it's an excellent example, because there's a large overlap with the existing NMS/EMS space. We started with RIB, then moved into BGP ... the higher we go in services, then we'll start to look more like an EMS/NMS. If this enables a new architecture for a more flexible NMS/EMS that can be enhanced more easily, then maybe that's still good.
Alex - General problem domain looks like EMS/NMS, so not sure the value-add from there. Pick use cases that are more narrowly defined.
Alia - Point of use case is not to standardize the use case, but the information that we need underneath the use case.
Jan - The question is why would you do that?
Himanshu - Larger point is you have a centralized place to collect information.
Jan - Trying to get us to think a layer higher.
Himanshu - This is a building block; once you have this, then you will build application.
Alia - Yes, agree. Its hard to see what the lines are setting dynamic state in the network vs. NMS/EMS state. Deals with speed, amount of change, etc. required to be learned from network and pushed into the network. Pulled these types of things from the existing Use Case drafts.
ChrisL - We can discuss if we want to do provisioning here or not. Other groups have already done this. One problem that has not been solved, with is a programmatic interface into the RIB. If we're looking for a space that hasn't been solved, it's programmatic interface to the RIB. Provisioning/configuration has already been solved.
Giles - For L3VPN, someone on the list has already asked: do you really need to use BGP to do that or could you use a controller to install the proper routes in the RIB. Perhaps we shouldn't have focused on L3VPN's, but should maybe thinking more about service-chaining. How do put that collection of devices into a service-chain.
Alia - (To Giles) don't think you did the wrong thing. You did a straightforward piece of it, this is the data we need back and events we need back. The question is: is that generating dynamic events back that you will react on vs. you can't do that today, e.g.: network events triggering certain events. Take a well-known use case and think beyond it, e.g.: what events could we react on, if we could get those events.
Luyuan - Existing L3VPN don't need something like this. But, svc-chaining in DataCenter, then things get much more dynamic, then value is probably more there.
Nitin - Not trying to rathole. In summary, you have to follow the money. No customer is stepping up to say that they want to use I2RS to rewrite their service provisioning mechanism.

Samita - Customers are saying that it takes weeks to add a service, but if we want to enable dynamic service provisioning then that would minimize the intervals to minutes. Subscriber services for mobile phone there are likely use cases.
Ed - Some technologies have come into play without “customer demand” not appearing beforehand.
Tom Narten (Jabber) - Do others see I2RS used for fast provisioning? Others are working on that, don’t see I2RS to do that.
Ed - Maybe others would use I2RS mechanisms to do that.
ChrisL - No argument about better automating provisioning, but we already have some of that.
Shane - Comment about bunch of things related to realizing a service (QoS, etc. attributes) vs. just manipulating routing information to forward packets on the network.
Alia - RIB, BGP and IGP are certainly in scope.
Tom Narten - +1 to Chris, think: VM’s + storage. This is largely a different area than I2RS.
Joel - If we try to do all of things related to service provisioning, then we’re biting off more than we can chew.
Jan - When we go higher up in the stack, then those info/data models are much harder to define.
Alia - We’re not trying to do services. If we try to do the whole space, then we’ll fail. We need to have a narrow set of vertically defined use cases. I need to identify the type of traffic, then I need to shove it into a tunnel and to do so I need to program the RIB to do that. We need narrowly defined use cases for that.

RIB -- Heart of I2RS -- Chris Liljenstople
- Scope - we never actually got to the read data portion of this
Not in scope - programming services and service-chaining; creating (dynamically or via config) and deleting RIBs
Routes programmed into RIB: unicast, multicast, MPLS
Routes could point to NH:
    - backup next-hop capability must be supported
    - backup next-hop may be combined with primary next-hop programming
Must be able to support multiple next-hops for a given route object
need a return code for: installed, active, reason (not authorized)
Route installed by I2RS client we need to know if its exportable to other routing protocols on the system
After route is installed: async notifications sent by I2RS agent to controller after route changes, e.g.: route isn’t installed or active any more
Scale - need Bulk API; instead of having one update per I2RS message, need to shove a bunch of routes in a single message. If we do bulk inserts, then we need to know which specific route(s) failed, vs. figuring out which one failed after-the-fact.
Pub/sub model is not requirement for events and notification -- don’t remember discussing that point
If controller inserted something stupid, then it’s problem of the controller
Debugging and troubleshooting is a separate problem?

Andy - Is there a need for transactions? I’m going to give you a bunch of things to do and you do them all or something like JSON patch.
Ed - Extremely valid question, but orthogonal from Chris’ question
ChrisL - Controller needs to stage the changes in the right order, so they get carried out in the proper order.
Andy - Had a conversation with Phil Shafer and really like their model that you can just throw things in there and the agent will sort it out. Adding xaction model on top of an edit-model is hard to do after-the-fact.
Alia - Talked about traffic destined for a particular prefix or destination, can you clarify that?
Chris - Talking about BGP & IGP variant, I believe. IOW, one is exit-point vs. next-hop.
Alia - Indirection of BGP and IGP. BGP says next-hop and IGP tells me how to get there.
Ed - The only way to avoid recursion is to point at the specific egress interface. That may be some way to

solve for it.
Ed - Used AFI/SAFI for specification of RIB rules. Are you implying we're going to specify flow specification rules.
ChrisL - Need rules for I2RS agent to tell the controller what AFI/SAFI can be installed into a FIB.
Andy - I'm assuming there's a broker and layered notification. Is there passthrough notifications through the broker to northbound applications.
Chris - I'm not sure we've gotten to that, yet.
Alia - Could be too early to discuss that.

Topology Use Cases -- Joel Halpern
Will work on an I-D
Worked out an Abstract, ToC based on summary of discussions we've held over the last few days
Joel - Will have it by the end of May

AAA/Security Stuff -- Wes George
- Agreed that there needs to be a draft, but need more volunteers to help flesh out the contents of the draft
- Should have cycles in the next month to get something drafty done
- Andy volunteered to help
- Adrian: do you want a security clueful person or one found for you?
- Wes: Let's get the draft put together first and figure out who we need.
- Adrian - OK, if you need help, please reach out and ask an A.D.
- Adrian - Jurgen has been volunteered by the Ops AD to monitor what's going on from an Ops PoV.

Customer/Service Topology -- Giles
- Focus:
(1) on information on what information we can pull from the router to aid provisioning.
(2) see if we can find other use cases that need to react to more dynamic information. Example might be: BCDR, Service Chaining that needs to react to a resource failing, etc.
- Nabil will have drafts written by the end of April ;)

RIB -- Nitin
- Focused on information model
- Use cases still need to be addressed, without that the rest of the picture will not be clear

Ed - Think there was a useful set of putative use cases.
Alia - Russ White also had a useful RIB draft, as well.
Nitin - Agreed.
Ed - Going to work on draft for Information Model.
Nitin - Yes
Ed - Use Cases?
Nitin - Will ask the team ... willing to co-author with ChrisL
Nitin - Timeframe: end of May.
Ed - Info model can be later.

BGP -- Didn't get a chance to discuss. No result. Ask again later.

Jan - Should we come up with a useful format for the contents of the informational model, e.g.: UML? Rest of group had problems with UML.

Ed - Problems with UML for Information Models
Joel - Topo Info Model that we're working on, UML could be more of an obstruction than a help, because you can't easily mark up annonations on edges of lines. Want to talk more about objects, their relationships to each other, etc.
Alia - Not fixated on UML
Joel - Want to focus first on describing information and then worry about getting it standard with others Info models later.
Alia - RIB folks were also looking at doing in UML
Nitin - Let first draft come out in whatever format and then we can morph it into whatever we need
Consensus in the room with the above point.

Wes - How are we going to take these discussions back to the larger WG and those drafts that were discussed today, but whose authors weren't in these discussions.
Ed - Suggest you reach out to those draft authors and figure it out.
Tom Narten (Jabber) - draft-ward-i2rs-framework: talked about BGP and BGP into the RIB, but didn't hear a lot of that talked about here. Is BGP still in scope?
Alia - BGP is definitely in scope, esp. because BGP is a large portion of the config. There could also be good use cases in the IGP, as well. Changing metrics based on events, spinning up VM's, etc. But, don't have any use cases written down for the IGP. Perhaps we look at re-chartering we'll pull IGP in scope.
Wes - How are we going to whittle down content? In particular, we talked about a lot of things here that aren't on the IETF 86 diagram of deliverables.
Ed - Slide shows diagram of what we talked about at IETF 86. This is what we're on the hook to deliver, but don't have thoughts on how we whittle that down, yet. Open to input on that.
Alia - Open to input on that. Russ' draft has a lot of the use cases, has a lot of the bones/content so hopefully that can be used for RIB Use Cases going fwd. Don't have an opinion on the others.
Ed - Don't think a single use case document for all use cases.
Wes - Agree with that. Wonder how you structure the use cases into draft.
Alia - So far, we've structured them into BGP and RIB and that's a good start for those. As others show up, we can figure it out.
Wes - As co-chairs, it may be helpful to others on the list what was discussed here today.
Tom Narten (Jabber) - Sounds like BGP in scope. IGP's later. All needs to be driven by use cases. When I say IGP I mean interacting with IGP directly, not via the RIB.
Alia - IGP is not in scope at the moment, but need to see use cases that would demonstrate a required change to the charter.
Himanshu - Is LSDB useful for Topology Info.?
Alia - Yes, it is. But, that's read-only view of Topology. What's not in scope is changing configuration of IGP..
Shane - the ability to inject summarisation routes somewhere like ASBRs would be useful. Not a RIB action. Wouldn't lose sleep if ruled out but seems easy to do. Nothing magical - normal practice today. Or what if I wanted to change the MD5 keys for IGP.
Alia - we're not chartered for an IGP information model or use cases but can talk about it when we recharter. But need to get stuff that's already in-charter done.
Himanshu - I took it to understand that I2RS interface is interface to read or write to the router. Need an interface to get information from the router. surely that's in-scope? What we really need to focus on is what objects do we need read or read-write access to and then figure out the actions we need and the protocol we want to use.
Alia - what we need to do now is have architecture and use-cases (ready for WG last call by Berlin as supposed to send to the IESG by August). Then have to do information models (bit more time there). From

use-cases and information model should get to protocol requirements and thence to reality rather than paper and slideware.
Shane - Himanshu raises a good point as to the confusion when we say e.g. “IGP is out of scope”. But we’ve clearly talked about e.g BGP-LS as a way to carry IGP information over BGP to a topology manager. is BGP-LS in scope of I2RS? No, but... Or a passive IGP listener figuring out the network topology. Critical requirement for I2RS. But different to changing the IGP configuration or injecting routes into the IGP. Those 2 use cases are out of scope. does that make more sense?
Tom - 3 types of access are worth distinguishing. No comment on which ones are in or out-of-scope at this time.

Ed - meeting now closed.