Network Scalability

Ignas Bagdonas
RIPE68
13-MAY-2014

Context

A BoF on L3VPN scalability was planned for IETF89 but was not held due to perceived lack of interest and problem space that is not aligned with IETF work.

IETF community might not be the right one for discussing such questions

Therefore a BoF on a broader and more generalized topic is being held here in RIPE to sense the interest, validate the problem space, and feed back the outcome to IETF

Scale BoF at IETF89

- **Non-WG forming BoF**
- **Initiative from MPLS and CCAMP WG chairs**
- **Focus on L3VPN scalability aspects**
- **Besides the L3VPN topics there were a few drafts on MPLS global label allocation and large labels on the agenda – and that resulted in push back from the community**
- **Proposed charter of the work perceived as too narrow, and if it is a L3VPN problem then it should be dealt with in L3VPN WG**
- **Multiple attempts by Routing ADs to initiate working discussions and eventually declared as lack of interest**

Scale BoF at IETF89

No BoF at IETF89

Topic is not closed, discussions endorsed to continue (*scale@ietf.org*)

Operator feedback needed – and IETF might not be the right community

Discussed with Routing and Operations ADs, Scale BoF chairs, several interested BoF members

Common suggestion was to align more with operator community

Scalability Context

Scalability of a system is a function of component properties and interactions of components that build the system

Amount of state

Rate of state change

Accumulation and distribution of state in different places in the network

Partial and full visibility of state

State

Entity representation in the network system

Elements in the system – nodes, interfaces, instances of logical constructs, messaging between entities.

Identifiers, addresses, data plane constraints

Prefixes, attributes, dependencies, inter-layer resolving

Practical Scalability

Control plane protocol mechanics

Control plane implementation specifics

Mapping to data plane

Data plane restrictions and implementation specifics

OAM, Telemetry

Architecture and desing influence to scaling properties

System level implementation specifics

Scope

Control plane protocol mechanics

- **In scope but likely little to cover here**

Control plane implementation specifics

- **In scope, possibly touching vendor specific aspects too**

Mapping to data plane

- **In scope with some uncertain areas**

Data plane performance

- **Not in scope, covered elsewhere (BMWG, LMAP) or considered too proprietary**

OAM, Telemetry

- **In scope as long as it does not go too close to data plane specifics**

Arhitecture and design influence to scaling properties

- **In scope and likely the main focus area**

System level implementation specifics

- **Possibly in scope**

Multicast Transport

Just blindly replicate at ingress ☺

Build and maintain trees for distribution

Track the interested receivers

Choice of signaling protocol and tunneling mechanism

Traffic Engineering

Just build a tunnel and shift the problem elsewhere ☺

Protection will require state to be preallocated

Midpoint state typically grows above linear compared to endpoint state

Tail end state may grow faster than head end

The choice of tunneling mechanism and signaling protocol

Composite Link OAM

Continuity Check in composite link environment may ensure that only the OAM path itself is verified to be continuous

Deterministic end to end continuity checking thus requires impractical amount of segment OAM state

It is easier to solve this problem in tunneled environment than in tunnel-less

MPLS Entropy labels

Stateless BGP and transport congestion

Platform implementations of traffic distribution

Platform dataplane capability export needed?

L3VPN

The "size" of VPN

Different visibility and different amount of state in different places in the network

Protection in L3VPN environment

Label ranges and label exhaustion

Signaling of intended and actual capacity across all network elements

Overlays

Just set up more tunnels ☺

Another layer and namespace of identifiers

Overlay and underlay OAM interaction

Overlay and underlay topology awareness

Restoration or topology change impact to overlay

Signalling between overlay and underlay

Non-routing Information in Routing Protocols

BGP the Universal Transport Protocol ☺

TE information flooding, topology information export

BGP-LS, INFORMATIONAL, address families used for signalling

Interrelation between information elements (related TLVs not fitting into the same protocol message)

Multidimensional metrics and metric components

Mapping Systems

Resolving identifier values of interrelated layers

Amount of outstanding requests

Impact to data plane operation (buffering is not a good idea)

Interworking with IETF

Problem statement and analysis done outside of IETF

Pain points identified

Requirements formulated for existing IETF WGs to address

Protocol work gets done in IETF

Way Forward

Is there a problem to be solved?

Are there any deployment experiences and best practices to be documented?

Where does it fit?

A subtopic in BCOP TF?

Discussion

Let's talk