NVO3 Architecture
draft-narten-nvo3-arch-01.txt
IETF88 – Vancouver
November 4, 2013

David Black, Jon Hudson, Larry Kreeger, Marc Lasserre, Thomas Narten
Changes -00 -> -01

• New section on distributed gateways
  – Optimize inter-VN communication, so NVEs tunnel directly to each other

• Replaced section on “push vs. pull”
  – Let’s not talk about “push” vs. “pull” generically
  – Interesting question are events leading to the need to obtain or propagate updates

• Improved text to better support adaptor offloads
NVE – NVA interaction

• For fault tolerance reasons:
  – Local NVA will be built out of multiple components
  – Individual components will have their own IP addresses

• How should NVE peer with local NVA?
  – Restricting NVA to single “floating” IP address is too limiting
  – Recommendation: require NVE support multiple addresses per NVA
    • NVE can failover to alternate addresses should NVA become unresponsive
    • Associate priorities with addresses to support load balancing, failover, etc.

• Followup question
  – Is all information available through a single address? Or can individual local NVA’s hold subsets of information (e.g., for particular VN’s)
  – Would simplify architecture if all information is available from any one peer
NVE – NVE interaction

- Data plane setup may involve NVE-NVE signaling
  - Security example: use IKE to set up IPsec between NVEs
- For control plane, if NVEs query other NVEs, why have NVA?
  - Requires full interconnection among NVEs
- However, NVEs can still provide hints related to forwarding:
  - No such VM at this location (but no indication of where it is)
  - VM has moved to another location (with pointer to new location)
  - But NVE should still check with NVA for authoritative answer
- Tenant multicast case tricky
  - When the membership and location of target VMs change, how can sender know?
  - Architecturally simpler if NVE can rely on NVA for this information
Data Plane Encapsulations

• There does not appear to be a need for a new encapsulation for NVO3
  – Key requirement is Context ID of sufficient size
  – Existing encapsulations seem adequate in practice
  – A place is needed, however, for maintenance and extension of “homeless” encapsulations (VXLAN, NVGRE)

• Control plane architecture should:
  – Support multiple encapsulations
  – Support gateways when NVEs do not share common encapsulation

• WG needs to make a decision one way or another
Next Steps

• WG needs to adopt and agree on what the architecture is
• Requirements difficult to specify when basic architectural choices have not yet been clearly decided
• Request formal adoption of document by WG
Questions?