IETF ARMD

Work Plan

Benson Schliesser, Linda Dunbar

IETF-82, Taipei

15 November 2011

Contents

- Current Status
  - Milestone Review
  - Existing Documents
- Next Steps

Milestones

May 2011	Problem statement
Nov 2011	ARP/ND statistics collection and behavior analysis in various Data Center environments
Nov 2011	Survey of Existing Implementations
Nov 2011	Survey of Security
Mar 2012	Recommendations to avoid or minimize issues caused by ARP/ND
Mar 2012	Gap Analysis

Existing Documents

- ARMD Problem Statement
  - http://tools.ietf.org/html/draft-ietf-armd-problem-statement-00
- Address Resolution Statistics
  - http://tools.ietf.org/html/draft-karir-armd-statistics-01
- DC Reference Architecture
  - http://tools.ietf.org/html/draft-armd-datacenter-reference-arch-01

Work Plan

1. Finish Existing Drafts
2. Develop Recommendations (BCP)
3. Examine Gaps, Document Requirements
4. Re-charter or Shutdown

Milestone Candidates

May 2011	Problem statement *(draft-ietf-armd-problem-statement-00,* *maybe draft-armd-datacenter-reference-arch)*
Nov 2011	ARP/ND statistics collection and behavior analysis in various Data Center environments *(draft-karir-armd-statistics-01)*
Nov 2011	Survey of Existing Implementations *(mentioned in draft-karir-armd-statistics,* *network design described in draft-armd-datacenter-reference-arch)*
Nov 2011	Survey of Security *()*
Mar 2012	Recommendations to avoid or minimize issues caused by ARP/ND
Mar 2012	Gap Analysis

Possible Additions to Existing Drafts

(For Discussion)

- Problem Statement
  - Host ARP/ND cache scaling?
  - Gratuitous ARP vs. Unsolicited ND?
  - 6MAN work on anti-DoS ND Improvements?
  - Security Survey?
- Statistics
  - Survey additional router platforms for Impact?
  - Host implementation survey?
  - Dual-stack Impact?
- Datacenter Reference
  - Leaf/Spine & Other Topologies?

Possible Recommendations

- ARP and/or ND Implementation
  - Host, Router
  - Timers, Security, …
- Operational "How to Scale"
  - Static Resolution
  - Limited Resolution Domains
    - Via L2/L3 Segmentation
    - Via Proxy Deployment

Gap Analysis

- What are our requirements (based on the problem statement)?
- What are the available solutions and/or our BCP recommendations?
- Is there a “gap” between these these?

Possible Gaps in our Gap Analysis

A. Documentation of ARP/ND Proxy
B. Overlay Inter-Layer Resolution
C. Others?

3. Examine Gaps, Document Requirements

A. Documentation of ARP/ND Proxy

ARP/ND Proxy Definition

(for reference)

- ND Proxy Behavior
  - RFC 4389
- ARP Proxy Behavior
  - Traditional
    - RFC 925, RFC 1027
  - Caching (“transparent”)
    - draft-shah-armd-arp-reduction
    - draft-sajassi-l2vpn-pbb-evpn-03
  - Mapping Directory and/or Distribution
    - draft-hu-trill-rbridge-esadi
    - draft-dunbar-trill-directory-assisted-edge
    - draft-xu-virtual-subnet-06
    - LISP
    - OpenFlow etc.
  - Other Variations

ARP Proxy Operational Requirements

- Proxy-to-Host and Inter-Proxy Behavior
- Timers
- Liveness Detection
- Effect on L2 Scaling
  - MAC learning etc?
- Effect on Host Implementation
  - Size of cache? Message rate?
- Others?

Overlay Inter-Layer Resolution

(for reference)

- Does the choice of overlay mechanism affect address resolution?
  - E.g. does the overlay control plane dictate how ARP proxy works?
- See for example
  - VPN4DC
    - Requirements for Seamless Multi-tenancy
  - NVO3
    - Overlay Architecture
    - Also see Loc/ID Split, etc
  - TRILL, VXLAN, NVGRE
    - Frame Format
    - MAC-in-MAC, MAC-in-IP, IP-in-IP, etc.
  - ESADI, SDNP, OpenFlow
    - Binding Distribution & Control Protocol

Req. for Overlay Address Resolution

- Overlay provides multi-tenant segmentation
- "Within" an Overlay
  - Same Problem Scope as ARMD
  - Except, overlay may define Proxy Behavior etc.
- "Outside" an Overlay
  - Binding of Inner Address to Outer Address

Others?

Backup Slides

Charter Highlights

- Context
  - significant increase in the number of hosts attached to the network
  - broadcast domains are scaling up to span more switching devices and VM servers, and to interconnect more hosts
  - heavily used protocols that are based on broadcast or multicast, such as ARP and ND, may contribute to poor network performance
- Our Aim
  - investigate the impact of changing workloads and existing protocols on datacenter network performance
- Objectives
  - (1) Document the current practices in data center network architectures and the scaling characteristics of ARP and ND with respect to large sized layer-2 domains in data centers
  - (2) Provide operational recommendations intended to minimize issues associated with these architectures and characteristics

Charter: WG Structure

- Area Affiliation
  - The armd Working Group is assigned to the Operations and Management area, and will maintain close collaboration with the Internet area
  - armd Working Group will focus on documenting current practices and scaling characteristics, and will not do any protocol development or extension work
- Way Forward
  - If the Working Group identifies opportunities for protocol development or extensions, it will first develop requirements for that work.
  - Any protocol development work will be conducted in the appropriate existing Working Groups if such work groups exist.
  - If no such working groups exist, armd may recharter to address the work and may be moved to a different area.

Charter: Deliverables

- Problem statement and review of current L2/ L3 architectures
- Report on ARP/ND statistics collection and behavior analysis in various Data Center environments
- Recommendations on data center L2/L3 architectures and identification of opportunities for protocol development work

http://tools.ietf.org/html/draft-ietf-armd-call-for-investigation-00

Call for Investigation

1. What are the scaling characteristics of modern datacenter networks (e.g. "dimensions" of scale and their normal ranges) that are relevant to address resolution?
2. What are the operational problems related to address resolution in the modern datacenter environment?
3. What is the relationship between scaling characteristics of datacenter networks (question #1) and operational problems related to address resolution (question #2)?
4. What, if any, are alternative solutions to the operational problems of address resolution at massive scale?
5. What, if any, are the "gaps" in existing solutions?

http://tools.ietf.org/html/draft-ietf-armd-problem-statement-00

ARMD Problem Statement

- Context
  - More endpoints, larger scope of L2 networks
  - Endpoint location decoupled from infrastructure
- Problems
  - Endpoint mobility limited by Address Resolution
  - Router (L3 gateway) message processing scale is limited
  - No ARP implementation standard (timers, retransmit, etc)
  - Packets are buffered pending address resolution
  - Learning Switch resources consumed by address resolution messages (via broadcast or multicast)

http://tools.ietf.org/html/draft-karir-armd-statistics-01

Address Resolution Statistics

- Factors that impact ARP/ND performance
  - Number of hosts
  - Traffic patterns
  - Network events
  - Implementations
  - L2 Topology
- Experiment Results
  - Message Rate grows (~linear) with number of hosts, traffic patterns; spikes with network events
  - Message Rate drives increased Router CPU utilization

http://tools.ietf.org/html/draft-armd-datacenter-reference-arch-01

DC Reference Architecture

- Defines an Access + Aggregation + Core 3-tier design taxonomy
  - L2/L3 edge can be at any tier
    - Impacts scope of VM mobility, scale of L3 gateway
  - Mentions Overlay approach to segmentation
- Offers Design Considerations
  - Traffic patterns, virtualization, etc
- Discusses L2/L3 boundary pain-point
  - ARP message scale impact
  - Different paradigm for ND unsolicited advertisements, versus gratuitous ARP