Multicast Overlay
Models & Mechanisms

Orlando IETF
March 2013

Dino Farinacci
farinacci@gmail.com
Agenda - Multicast Delivery Models

- Multicast-over-Unicast (MBONE)
- Multicast Native
- Virtualizing Multicast - MVPNs
- Multicast-over-Unicast (AMT)
- Multicast Map-and-Encap Overlays
Agenda - Multicast Overlay Signaling Mechanisms

• In-the-Network Signaling
  • RFC 6831 - LISP-Multicast (PIM)
  • draft-farinacci-lisp-mr-signaling (LISP)

• Out-of-the-Network Signaling
  • Mapping Database Based (PIM and/or LISP)
    • draft-coras-lisp-re & draft-ietf-lisp-lcaf
  • Programmable Interface
    • i2rs, OpenFlow, RESTful
Native Arrived (kind of)
ISP Multicast Service

Customer 1

Customer 2

PIM multicast capable

within-ISP multicast capable

MVPN
We Wanted Multicast Anywhere
Now We Have Overlays

- unicast/multicast capable
- unicast ubiquitous multicast disjoint
- unicast/multicast capable
- unicast/multicast capable
- unicast/multicast capable

EIDs -> green
RLOCs -> red
LISP-Multicast Today

- RFC 6831 - draft-ietf-lisp-multicast
  - Defines how to encap multicast into multicast or unicast
  - Defines use of unicast PIM J/P message exchange between ETRs and ITRs
  - Defines how to work with native PIM at source and receiver multicast sites
  - Enumerates various combinations and recommends how to avoid combinatoric nightmares
Core Supports Native Multicast

- (S-EID, G)-multicast capable
- (RLOC, DG)
- No state
- (EID, G) tree state
- Link with no tree state
- (RLOC, DG) tree state

- LISP xTRs
- Hosts in EID space
- Routers in EID space
- Routers in RLOC space
- Replicating router
Multicast Overlay
Signaling Mechanisms
In-the-Network Signaling

• Use traditional protocol based signaling methods?
  • RFC 6831 - using PIM
  • draft-farinacci-lisp-mr-signaling - using LISP
PIM Control-Plane Everywhere

 multicast capable

No state
LISP-Multicast Tomorrow

- Eliminate the need for PIM over-the-top
  - Less protocols mean lower OpEx and less complexity
- Use the existing mapping system for ETRs to find ITRs of source multicast sites
- At the same time allow for encap of multicast into unicast
  - To allow multicast service over partner unicast-only network
LISP as Control-Plane

multicast capable

Map-Request

Map-Reply

LISP

PIM

S

R

Map-Request

Map-Reply

LISP

multicast capable

Map-Reply

LISP

PIM

R

PIM

R

PIM

PIM
Out-of-the-Network Signaling

- Use the Mapping Database
  - Replication list of ETRs or DGs are stored per (S-prefix, G-prefix) EID entry
  - See LISP Replication Engineering (LISP-RE) design

- Use a Programmable Interface
  - Have network controller monitor ETRs for joined state
  - Then network controller programs ITRs with replication state
  - Network controllers can program RTRs inside of network to optimize distribution trees
Using the Mapping Database

(S-EID, G) encoding

Multicast Info Canonical Address Format:

```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| AFI = 16387 | Rsvd1 | Flags |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Type = 9 | Rsvd2 | R|L|J | 4 + n |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Reserved | Source MaskLen | Group MaskLen |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| AFI = x | Source/Subnet Address ... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| AFI = x | Group Address ... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

Replication List Entry Address Format:

```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| AFI = 16387 | Rsvd1 | Flags |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Type = 13 | Rsvd2 | 4 + n |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Rsvd3 | Rsvd4 | Level Value |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| AFI = x | RTR/ETR #1 ... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Rsvd3 | Rsvd4 | Level Value |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| AFI = x | RTR/ETR #n ... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

RLOC or DG encoding
Map-Cache entry in A:
(SEID, G) -> 
multicast DG {which B & C joined (RLOC-A, DG)}
unicast RLOC D
unicast RLOC E
Programmability Example

Network controller reads from:
- RLOC D wants (SEID, G) via unicast
- RLOC E wants (SEID, G) via unicast
- RLOC B wants (SEID, G) via DG
- RLOC C wants (SEID, G) via DG

Network controller writes to:
- RLOC B to join (RLOC-A, DG)
- RLOC C to join (RLOC-A, DG)

Network controller writes to:
- RLOC A to add (SEID, G) entry with:
  - multicast DG
  - unicast RLOC D
  - unicast RLOC E

unicast  multicast
Advanced Topic - Future

- If unicast replication becomes popular ...
  - ... need to manage head-end replication overhead
- Will need in-the-network replicators (like AMT Relays)
- See draft-coras-lisp-re-02
Q&A

Multicast can turn any simple problem into a hard one

:-)