OPSAWG Working Group
INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: September 2012

Balaji Venkat Venkataswami
Bhargav Bhikkaji
DELL-Force10
March 26, 2012

VM to VTEP maps topology discovery in VXLAN based data centers
draft-balaji-opsawg-vxlan-vm-topo-discovery-01

Abstract

This document proposes a method by which in a VXLAN environment the ARP tables of each VTEP having an active VM belonging to a particular tenant where such active VMs are distributed amongst several VTEPs in a data center or across data centers are walked through and the collation of the location of such active VMs and the VTEPs they are located in is found for management and network resource planning purposes.

Table of Contents

1 Introduction

It will be necessary in a VXLAN data center environment to locate the several active VMs belonging to one or more tenants or all tenants which are hosted by the VTEPs in the data center and list the active VMs such that management and network resource planning can be done for that tenant. This information may be useful to the network administrators of the data center deploying VXLAN and to the tenants that have their active VMs hosted in the data center running VXLAN for the mentioned purposes.

1.1 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

1.2 Methodology

The following IP Address Translation table also called the ARP table as outlined in RFC 2011 will be useful in this method. It is possible that several IP sub-nets are deployed for a given tenant. For each such IP sub-net there may be a VLAN allocated. For routing between such IP sub-nets the default gateway for a VLAN that has its VNICs in several such VLANs may be deployed. In such a case the ARP tables of each of the interfaces of then active VM default gateway for the several VLANs involved is inspected to collate the different VLAN's active VMs that are hosted on the VTEPs in the VXLAN based data center. The algorithm that does this specific job of collation is explained in section 1.2.1.

```
-- the IP Address Translation table

-- The Address Translation tables contain the IpAddress to
-- "physical" address equivalences.  Some interfaces do not
-- use translation tables for determining address
-- equivalences (e.g., DDN-X.25 has an algorithmic method);
-- if all interfaces are of this type, then the Address
-- Translation table is empty, i.e., has zero entries.

ipNetToMediaTable OBJECT-TYPE
    SYNTAX      SEQUENCE OF IpNetToMediaEntry
    MAX-ACCESS  not-accessible
    STATUS      current
```

```
        DESCRIPTION
                "The IP Address Translation table used for mapping from
                IP addresses to physical addresses."
        ::= { ip 22 }

    ipNetToMediaEntry OBJECT-TYPE
        SYNTAX      IpNetToMediaEntry
        MAX-ACCESS  not-accessible
        STATUS      current
        DESCRIPTION
                "Each entry contains one IpAddress to `physical' address
                equivalence."
        INDEX       { ipNetToMediaIfIndex,
                      ipNetToMediaNetAddress }
        ::= { ipNetToMediaTable 1 }

    IpNetToMediaEntry ::= SEQUENCE {
            ipNetToMediaIfIndex      INTEGER,
            ipNetToMediaPhysAddress  PhysAddress,
            ipNetToMediaNetAddress   IpAddress,
            ipNetToMediaType         INTEGER
        }

    ipNetToMediaIfIndex OBJECT-TYPE
        SYNTAX      INTEGER (1..2147483647)
        MAX-ACCESS  read-create
        STATUS      current
        DESCRIPTION
                "The interface on which this entry's equivalence is
                effective.The interface identified by a particular value
                of this index is the same interface as identified by the
                same value of RFC 1573's ifIndex."
        ::= { ipNetToMediaEntry 1 }

    ipNetToMediaPhysAddress OBJECT-TYPE
        SYNTAX      PhysAddress
        MAX-ACCESS  read-create
        STATUS      current
        DESCRIPTION
                "The media-dependent `physical' address."
        ::= { ipNetToMediaEntry 2 }

    ipNetToMediaNetAddress OBJECT-TYPE
        SYNTAX      IpAddress
        MAX-ACCESS  read-create
        STATUS      current
        DESCRIPTION
                "The IpAddress corresponding to the media-dependent
```

```
                `physical' address."
        ::= { ipNetToMediaEntry 3 }

    ipNetToMediaType OBJECT-TYPE
        SYNTAX      INTEGER {
                    other(1),        -- none of the following
                    invalid(2),      -- an invalidated mapping
                    dynamic(3),
                    static(4)
                }
        MAX-ACCESS  read-create
        STATUS      current
        DESCRIPTION
                "The type of mapping.
                Setting this object to the value invalid(2) has the
    effect
                of invalidating the corresponding entry in the
                ipNetToMediaTable.  That is, it effectively disassociates
                the interface identified with said entry from the mapping
                identified with said entry.  It is an implementation-
                specific matter as to whether the agent removes an
                invalidated entry from the table.Accordingly, management
                stations must be prepared to receive tabular information
                from agents that corresponds to entries not currently in
                use.  Proper interpretation of such entries requires
                examination of the relevant ipNetToMediaType object."
        ::= { ipNetToMediaEntry 4 }
```

1.2.1 Algorithm

```
Input : Seed VTEP IP address of a particular tenant Y

Output: Collated output of all active VMs in the respective VTEPs in
the VXLAN data center.

AlgorithmBegin

While more VTEPs to be scanned

START_LABEL:

     While ( there exists more entries in
     current_vlan ARP TABLE Where the TABLE = RFC 2011
     ipNetToMediaEntryTable of Seed VTEP )

             Get Next of the entry in the ARP table of the VTEP;
                     If (active VM listed in ARP table is
                     tenant of Y )
             then
                     Add to list the unique ARP table entry;
                     Add VTEP in the ARP table entry
                     to unique VTEP list;
             endif

     EndWhile

     If (any other VLAN's ARP table is available
             in case the VM is a gateway VM) then

             Set current_vlan = VLAN located;
             goto START_LABEL;
     else
             // do nothing;
     endif

     Set Seed VTEP = Next VTEP address in the unique VTEP list;

     Advance one entry in the unique VTEP list;

     Set NextVTEP = Seed VTEP;

EndWhile;

AlgorithmEnd;
```

2. Applicability to NMS Applications

Network Management Applications can provide a friendly user interface where the topology of the Layer 3 transport network with the TORs and respective VTEPs under them can be discovered using regular Layer 3 topology discovery. The algorithm in 1.2.1 can then be executed and the active VMs of various tenants displayed. This will help in management and in network resource planning.

2.1 VTEP support

VTEPs in the VXLAN environment in data centers are expected to have SNMP support in the form of MIBs as per [2011].

3 Security Considerations

The usual SNMP related security concerns apply.

4 IANA Considerations

None.

5 References

5.1 Normative References

[KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC1776] Crocker, S., "The Address is the Message", RFC 1776, April 1 1995.

[TRUTHS] Callon, R., "The Twelve Networking Truths", RFC 1925, April 1 1996.

5.2 Informative References

[RFC2011] McCloghrie, K., Ed., "SNMPv2 Management Information Base for the Internet Protocol using SMIv2", RFC 2011, November 1996.

[EVILBIT] Bellovin, S., "The Security Flag in the IPv4 Header", RFC 3514, April 1 2003.

[RFC5513] Farrel, A., "IANA Considerations for Three Letter Acronyms", RFC 5513, April 1 2009.

[RFC5514] Vyncke, E., "IPv6 over Social Networks", RFC 5514, April 1 2009.

Authors' Addresses

Balaji Venkat Venkataswami,
Dell-Force10,
Olympia Technology Park,
Fortius block, 7th & 8th Floor,
Plot No. 1, SIDCO Industrial Estate,

Guindy, Chennai - 600032.
TamilNadu, India.
Tel: +91 (0) 44 4220 8400
Fax: +91 (0) 44 2836 2446

EMail: BALAJI_VENKAT_VENKAT@dell.com

Bhargav Bhikkaji,
Dell-Force10,
350 Holger Way,
San Jose, CA
U.S.A

Email: Bhargav_Bhikkaji@dell.com