A SAVI Solution for DHCP

Jun Bi, Jianping Wu, Guang Yao, Fred Baker
draft-ietf-savi-dhcp-01(02).txt
IETF77, Anaheim
Mar.23 2010
Outline

• Solution Basis
• Additional Features in 01(02) Version
• Next Step
Solution Basis
Basis and Related Protocols

• A control packet snooping based solution. Data packet snooping is used as supplement.

• **Stage 1: DHCP Address Assignment**
  – DHCPv4(RFC2131)
  – DHCPv6(RFC3315, stateful)

• **Stage 2: Duplicate Detection**
  – IPv4 Address Conflict Detection(RFC5227)
  – IPv6 Duplicate Address Detection(RFC4862)

• Optional Data Trigger function to handle some cases:
  – Will be discussed in 2nd part of this PPT
Typical Scenario

The Router or SAVI device may also play the role of DHCP Relay (or even DHCP server) in implementation.
# Anchor Attributes

- **Attribute**: Configurable features of anchor (anchor could be a port at a switch)
- An anchor may be configured to one or more compatible attributes, depending on the requirement of administrator

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAVI-Validation</td>
<td>Snooping &amp; Filtering</td>
</tr>
<tr>
<td>SAVI-SAVI</td>
<td>No binding and no filtering, trusted</td>
</tr>
<tr>
<td>SAVI-DHCP-Trust</td>
<td>Trust DHCP server type message</td>
</tr>
<tr>
<td>SAVI-LocalGroup(Optional)</td>
<td>Share binding entries at multiple anchors</td>
</tr>
<tr>
<td>SAVI-DataTrigger(Optional)</td>
<td>Allow data triggered binding process</td>
</tr>
</tbody>
</table>
SAVI-LocalGroup Attribute

• Handle the scenario that multiple anchors used by the same group of clients.
  – A group must be identified by name or index

Group 1: anchor 1, anchor 2

Or

Anchor 1:
  SAVI-LocalGroup group 1

Anchor 2:
  SAVI-LocalGroup group 1
SAVI-DataTrigger Attribute

- Handle special case
  - Local link movement
  - Link layer topo change or layer-2 path change

- Diagram showing movement of SAVI devices and hosts with DHCP server and data packets.
## Compatibility between Attributes

<table>
<thead>
<tr>
<th></th>
<th>SAVI-Validation</th>
<th>SAVI-SAVI</th>
<th>SAVI-DHCP-trust</th>
<th>SAVI-LocalGroup</th>
<th>SAVI-DataTrigger</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAVI-Validation</td>
<td>_</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>SAVI-SAVI</td>
<td>N</td>
<td>_</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>SAVI-DHCP-trust</td>
<td>Y</td>
<td>N</td>
<td>_</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>SAVI-LocalGroup</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>_</td>
<td>Y</td>
</tr>
<tr>
<td>SAVI-DataTrigger</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>_</td>
</tr>
</tbody>
</table>
Conceptual Data Structures

- **Control Plane: Binding State Table (BST)**
  - Keep **state** and **lifetime**
  - Key on anchor and(or) address
  - Entry: *Anchor | *Address | State | Lifetime | Other
- **Data Plane: Filtering Table (FT)**
  - Used for filtering only (for instance, ACL)
  - Key on anchor
  - Entry: *Anchor | Address
- **BST and FT can be combined or separated in implementation.**
Prefix Configuration

• Prefix scope can be learnt by
  – Automatically from RA or DHCP-PD
  – Manually configuration

• Optional configuration
  – entirely trust the DHCP server
States of binding

• **START**  A DHCP request (or a DHCPv6 Confirm, or DHCPv6 Solicitation with Rapid Commit option) is received from host, and it may trigger a new binding.

• **LIVE**  A DHCP address is acknowledged by a DHCP server.

• **DETECTION**  A gratuitous ARP or Duplicate Address Detection NSOL has been sent by the host (or **SAVI device**).

• **BOUND**  The address has passed duplicate detection and it is bound with the anchor.
State Transit Diagram

Red: from hosts, orange: from dhcp server

- Red: from hosts, orange: from dhcp server

Start

Live

Detection

Bound

Timeout

DHCP Request/Confirm

DHCP Reply

Response for Detection

Detection Packet

Timeout => Send Probe

Detection Timeout

Decline

Lease time expires/ DHCP release/ decline

Reply for DHCP renew/rebind
## State transit table

<table>
<thead>
<tr>
<th>State</th>
<th>Packet/Event</th>
<th>Action</th>
<th>Next State</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>Request/Confirm</td>
<td>Set up new entry</td>
<td>START</td>
</tr>
<tr>
<td>START</td>
<td>ACK/Reply</td>
<td>Record lease time</td>
<td>LIVE</td>
</tr>
<tr>
<td>*START</td>
<td>ACL/Reply</td>
<td>Recording lease time. Send probe</td>
<td>DETECTION</td>
</tr>
<tr>
<td>LIVE</td>
<td>DAD NS/Gratuitous ARP</td>
<td>-</td>
<td>DETECTION</td>
</tr>
<tr>
<td>LIVE</td>
<td>DECLINE</td>
<td>Remove entry</td>
<td>-</td>
</tr>
<tr>
<td>LIVE</td>
<td>Timeout</td>
<td>Send ARP Req/NS</td>
<td>DETECTION</td>
</tr>
<tr>
<td>DETECTION</td>
<td>Timeout</td>
<td>-</td>
<td>BOUND</td>
</tr>
<tr>
<td>DETECTION</td>
<td>ARP RESPONSE/NA</td>
<td>Remove entry</td>
<td>-</td>
</tr>
<tr>
<td>DETECTION</td>
<td>DECLINE/RELEASE</td>
<td>Remove entry</td>
<td>-</td>
</tr>
<tr>
<td>BOUND</td>
<td>RELEASE/DECLINE</td>
<td>Remove entry</td>
<td>-</td>
</tr>
<tr>
<td>BOUND</td>
<td>Timeout</td>
<td>Remove entry</td>
<td>-</td>
</tr>
<tr>
<td>BOUND</td>
<td>Reply on RENEW/REBIND</td>
<td>Set new lifetime</td>
<td>BOUND</td>
</tr>
</tbody>
</table>
Filtering Specification

• For anchor with SAVI-Validation attribute:
  – Data packet:
    • Check if <anchor, source address> in Filtering Table
  – Control packet (DHCP, NDP, ARP):
    • DHCPv4 Discovery: source address MUST be all zero
    • DHCPv4 Request: source address MUST be all zero or a bound address
    • DHCPv6 Request/Confirm: source address MUST be a bound address (either SLAAC or DHCP or manual)
    • DHCP Reply/Ack MUST be from port with SAVI-DHCP-Trust Attribute
    • NSol/ARP Request: source address MUST be a bound address (or unspecified address in case of DAD NS)
    • NAdv/ARP Reply: source address and target address MUST be bound addresses.
Binding Removal

• If the lifetime of a binding entry expires
• If the host is off-link
• If a local link movement is confirmed
  – Local link movement may be confirmed when address is assigned to another anchor and no conflict (DAD is successful)
Additional Features in 01(02) Version
Handle Anchor Off-Link Event

• If an anchor with SAVI-Validation is off-link
  – Keep the entry for a short period (for cable connection unstable case).
  – If the anchor turns on-link during the period, keep the bindings.
  – After the period, if it’s still off-link, delete the bindings.
Binding Number Limitation

• Avoid DoS exhausting the Binding State Table
  – Three choices
    – Set the upper bound of binding number for each anchor with SAVI-Validation.
    – Reserve a number of binding entries for each anchor with SAVI-Validation attribute and all anchors share a pool of the other binding entries.
    – Limit DHCP Request rate per anchor, using the bound entry number of each anchor as reverse indicator.
CONFIRM triggered binding

• CONFIRM message is replied with status of address but not lease time.
• The SAVI device should retrieve the lease time of the bound address using LEASEQUERY, if the address is not assigned, the binding should be removed.
State Restoration

• The SAVI device may lose binding states because of scheduled or unexpected reboot
  – If the switch directly connects to hosts, then bindings will be recovered by hosts
  – There were lots of discussions on mailing-list for remote switch reboot, we have 3 optional ways
  – The bindings should be stored into non-violate storage regularly or manually (proposed by Mikael)
  – Or upstream router send NS triggered by 801.ag then savi-device binds by NA (proposed by John)
  – Or use the optional data triggered probes during a short period after reboot (see next pages)
Data Trigger Procedure(1)

• Data trigger function is enabled on anchor with SAVI-DataTrigger attribute

• Whenever a packet whose source is not in the Filtering Table is received, the SAVI device:
  – Drop the packet in case the address is bound on another anchor.
  – Send a DHCP LEASEQUERY message, and wait for the result. (The data packet should be forwarded or discarded during the waiting time, but not stored)
Data Trigger Procedure(2)

• If SAVI device is pure layer 2 switch with no layer 3 address
  – If stateless address is also permitted (not dhcp-only)
    • Use DAD to check whether the address is being used by another anchor. Allow the address if no conflict.
    • Then rebind the address if DHCP renew/rebind is received.
  – If not
    • User recover binding by repair the network connections
    • Or configure a short DHCP lease time. Then user can repair binding automatically.
Next Step
Next Step

• Plan to submit savi-dhcp-02 officially based on feedbacks from IETF77
• Please provide comments during the meeting or in the mailing-list
• Solution had been implanted by multiple vendors and being deployed in CNGI-CERNET2 (will be reported in my next PPT)
• May consider to ask for last call in IETF78
Thank you very much!