A SAVI Solution for DHCP

Jun Bi
CERNET/Tsinghua Univ.
draft-jbi-savi-dhcp-00
IETF76, Hiroshima
Nov.9 2009

Outline

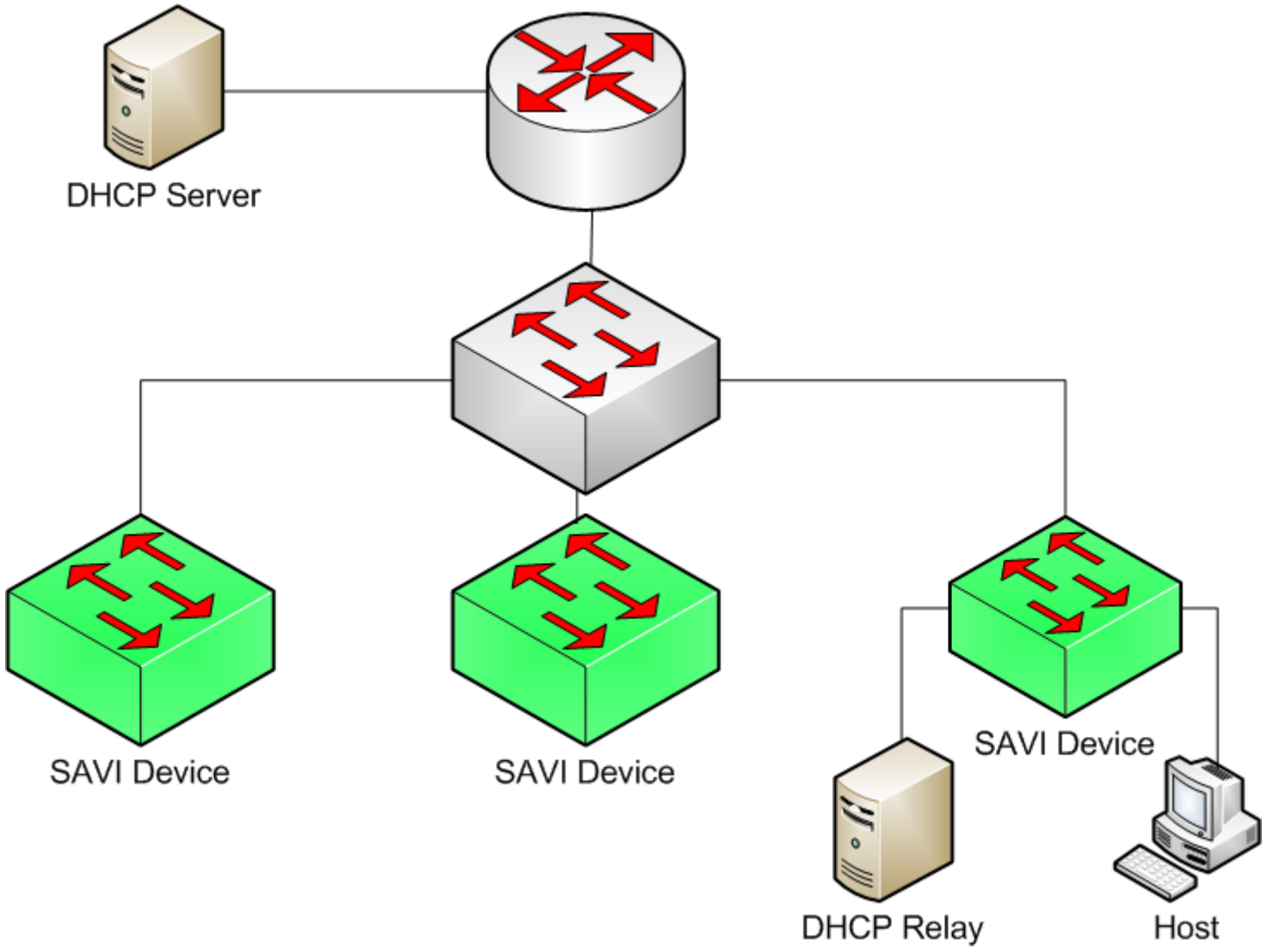
- Solution Overview
- Handling Special Situations
- Open Issues
- Implementation and Next Step

Solution Overview

Basis and Related Protocols

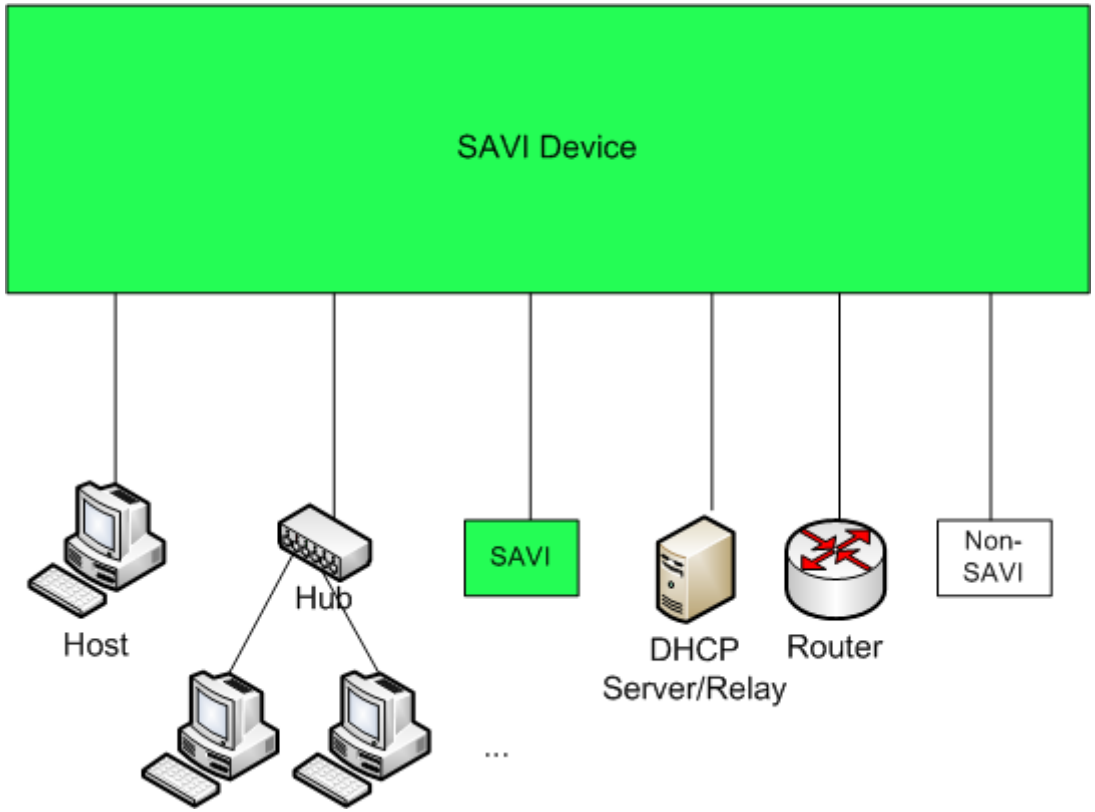
- A control packet snooping based solution. Data packet snooping is used as supplement.
- *Stage 1: DHCP Address Assignment*
  - DHCPv4(RFC2131)
  - DHCPv6(RFC3315, stateful)
- *Stage 2: Duplicate Detection*
  - IPv4 Address Conflict Detection(RFC5227)
  - IPv6 Duplicate Address Detection(RFC4862)

Typical Scenario



Port (Trust Anchor) Types

Might be moved to framework



Type	Action
SAVI-Host (Recommend to deploy)	Snooping & Filter(most secure and light-weight)
SAVI-Poly	Snooping & Filter
SAVI-SAVI	No binding and no filtering
SAVI-DHCP-Trust	Trust DHCP Reply
SAVI-nonSAVI	*Suggest to separate from SAVI area by VLAN*
SAVI-Router	*No define & no action*

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.

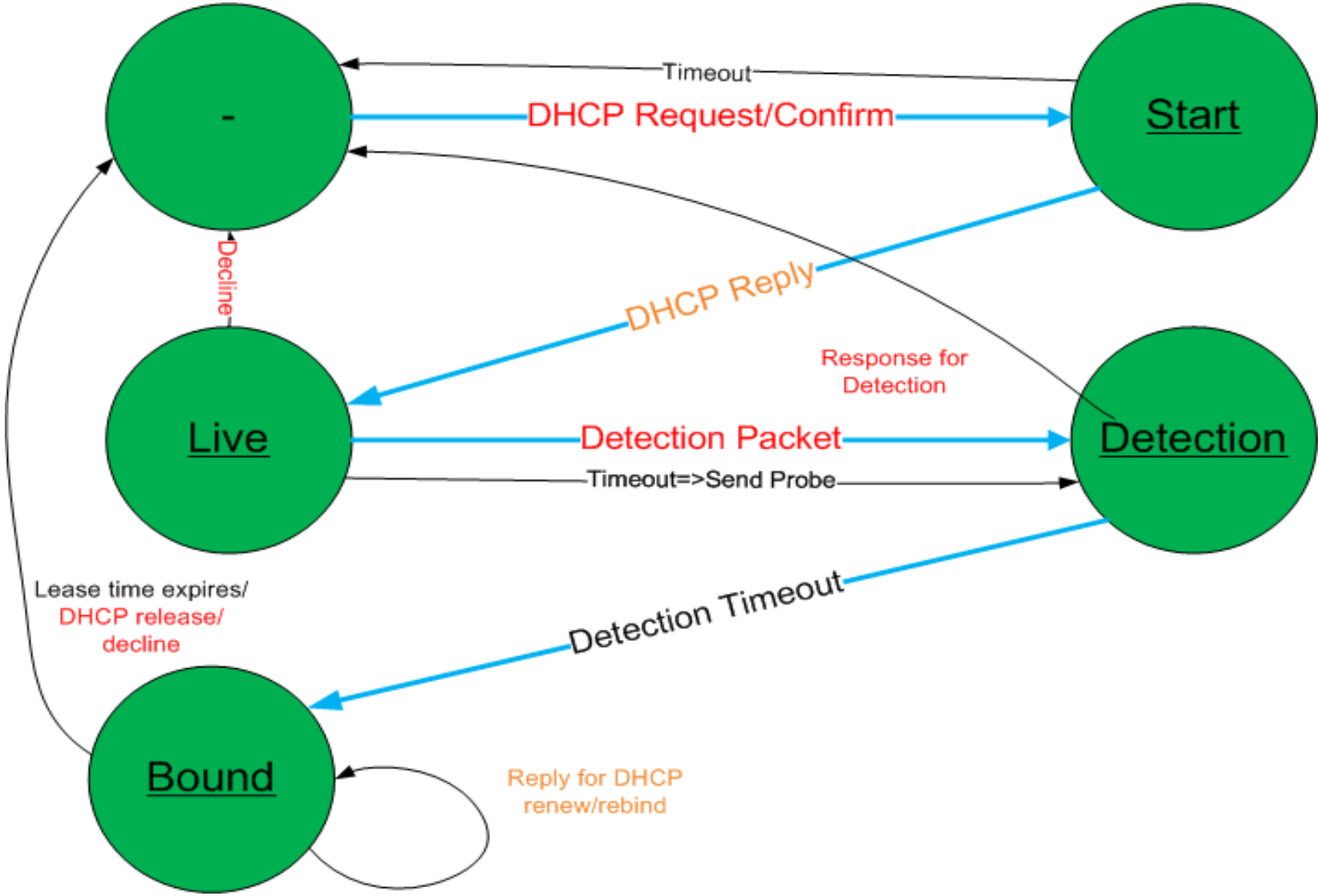
Prefix Configuration

- Configure reasonable prefix scope
  - Learn from RA or DHCP-PD
  - Manually configuration
- *Open issue*: Trust DHCP server or trust prefix configuration when DHCP acknowledged address is in conflict with the prefix?
  - Reason: Malicious/Fake DHCP server
  - If trust prefix configuration, then drop the malicious DHCP-reply
- *Security issue*: Keep RA secure
  - RA guard(draft-ietf-v6ops-ra-guard-03)?
  - Or SAVI-RA-Trust port for simplicity?
- Might be moved to framework

States of binding

- START A DHCP request (or a DHCPv6 Confirm) 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 Transition Diagram



State transition table

State	Packet/Event	Action	Next State
-	Request/Confirm	Set up new entry	START
START	ACK	Record lease time	LIVE
START	Timeout	Remove entry	-
LIVE	DAD NS/Gratuitous ARP	-	DETECTION
LIVE	DECLINE	Remove entry	-
LIVE	Timeout	Send ARP Req/NS	DETECTION
DETECTION	Timeout	-	BOUND
DETECTION	ARP RESPONSE/NA	Remove entry	-
DETECTION	DECLINE	Remove entry	-
BOUND	RELEASE/DECLINE	Remove entry	-
BOUND	Timeout	Remove entry	-
BOUND	Reply on RENEW/REBIND	Set new lifetime	BOUND

Filtering Specification

- Data packet:
  - Filter packet from SAVI-host and SAVI-poly port
  - By checking if (anchor, source) in Filtering Table
- Control packet(DHCP, NDP, ARP):
  - DHCPv4 Request/Discovery: source address MUST be all zero
  - DHCPv6 Request/Confirm: source address MUST be a bound address(either SLAAC or DHCP or manual, at least link-local)
  - DHCP Reply/Ack MUST be from port with SAVI-DHCP-Trust
  - 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 a bound address.

Binding Number Limitation

- Set a limitation per port to stop DoS against binding table.
- Or a adaptive rate limit mechanism with the similar effect.
  - Request rate limitation depends on current binding entry number on the port

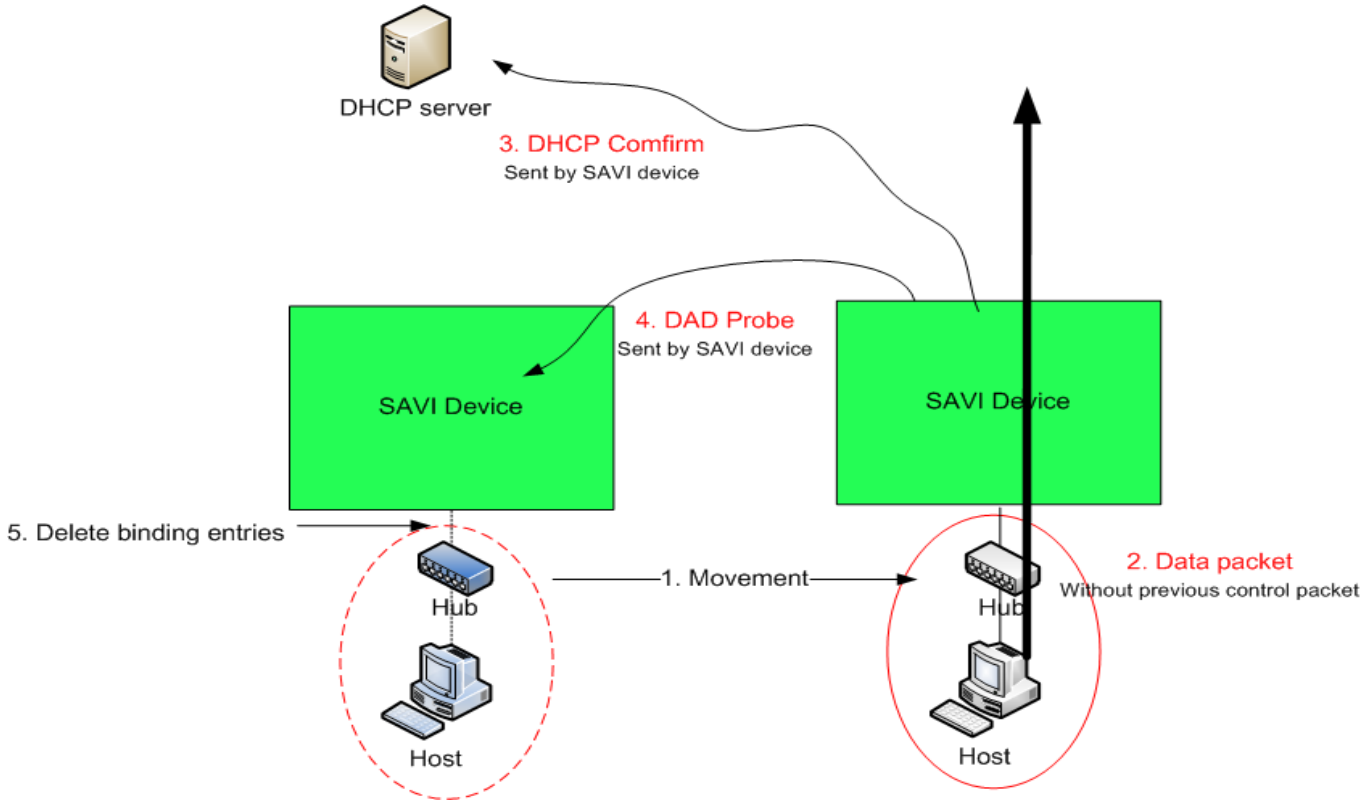
Handling Special Cases

Usage of Probe in Special Cases

- Usage of probe in special cases (will be explained in next slides)
  - Movement detection at poly-port
    - DAD/Gratuitous ARP
    - Not deliver to the source port
  - Alive detection: port down/up for assurance
    - NUD/ARP Request
  - Hold binding for inactive node
    - NA/ARP Response
- Format of probe
  - DAD/Gratuitous ARP: link layer address of Host
  - NUD/ARP Request: IP address and link layer address of SAVI device (switch management address)
  - NA/ARP Response: the link layer address and IP address of host

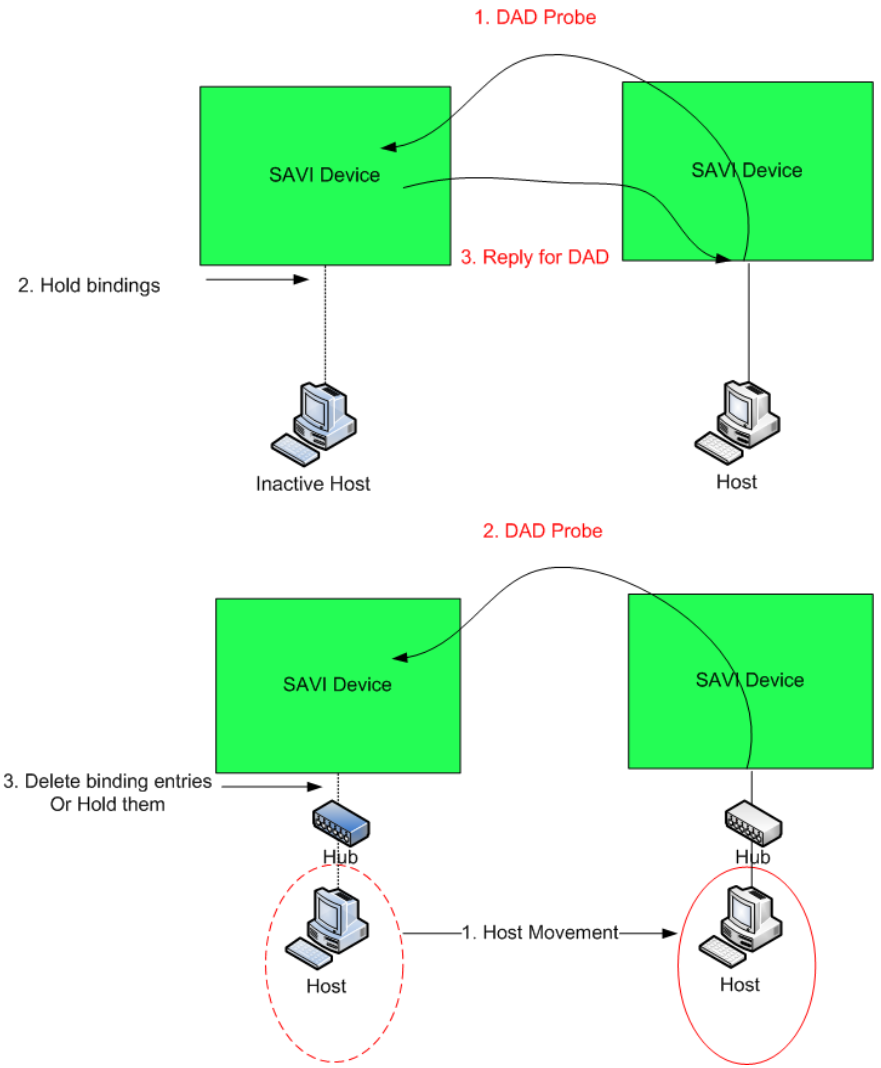
Data Packet Snooping at SAVI-Poly port

- Handle moving from one SAVI-Poly port to anther SAVI-Poly port.
- No DHCP procedure will be triggered at the host after moving!
  - Different from movement at SAVI-Host port (host sending DHCP-Confirm)
- A DHCP confirm will be sent by the SAVI device then a DAD probe will be triggered and the old binding will be removed



Binding Remove

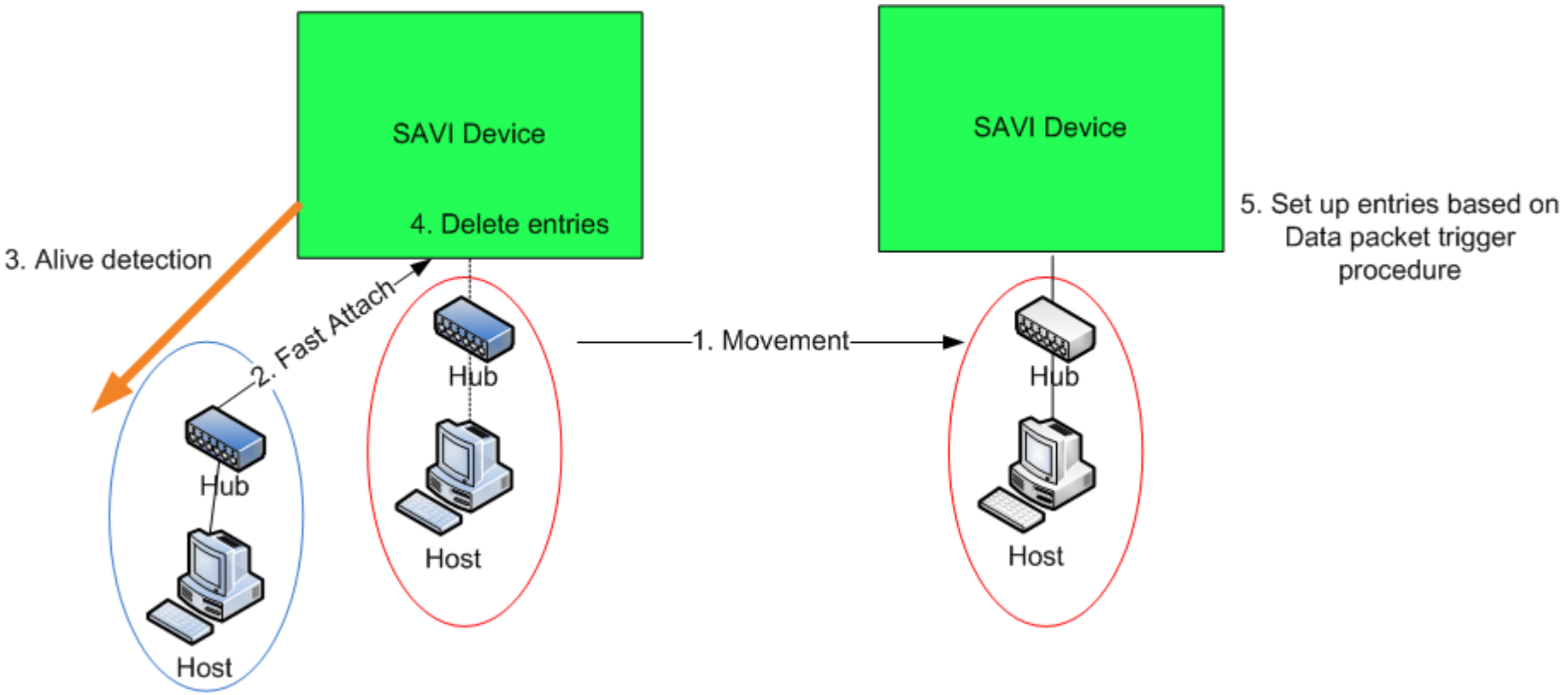
- MUST (Normal case):
  - Remove an binding entry whenever lifetime expires.
- MAY (Special case):
  - When the SAVI device receives a DAD NS/Gra ARP request target at an address bound and there is no reply from the port (the link is up)
  - At SAVI-host port, hold binding for Host (inactive node) by sending NA/ARP response
  - At SAVI-poly port, remove (for host movement) or hold (for inactive node)
- Other situations discussed in "Port down event"



Port Down Event

- SAVI-Host/SAVI-Poly port
  - To handle flappy links, keep binding entries of the port with link down event for a very short time. After the period, remove the entries.
  - To handle movement, if receiving DAD NS/Gra ARP request target at the address during the period, remove the entry.
  - If port turns UP during the period
    - Optionally send probes to SAVI-host port for assurance
    - MUST send probes to SAVI-Poly port for assurance (to handle a very special case, see next slide)

Port Down/Up Event at SAVI-Poly port



Open Issues

Open issues

- Whether to keep START state
  - Benefits:
    - Bind address and anchor securely (know exact source port of DHCP-request)
    - Limit Request rate to protect DHCP server
  - Defects: Temporary states (may be dangerous at SAVI-ploy port, but it's OK at SAVI-host)
  - Optional (contributed by Eric Levy-Abegnoli)
    - If MAC is unspoofable, then we don't need START state
    - Insert option 82 into packet
      - But not all servers support for option 82
      - Burden for SAVI switch to act as DHCP RELY

Implementation and Next Step

Implementation and Next Step

- Currently, this solution has been implemented by multiple vendors and is being deployed in Tsinghua Campus/CERNET2
  - will be reported in my next PPT (CNGI-CERNET SAVI deployment update)
- Can we move forward with this document as the basis of ietf-savi-dhcp-00

Thank you very much!

Q&A