Network Working Group
Request for Comments: 2281
Category: Informational

T. Li
Juniper Networks
B. Cole
Juniper Networks
P. Morton
Cisco Systems
D. Li
Cisco Systems

March 1998

Cisco Hot Standby Router Protocol (HSRP)

Status of this Memo

Copyright Notice

IESG Note

This document reflects an existing deployed protocol. The IETF does have a working group which is in the process of producing a standards track protocol to address the same issues.

Abstract

The memo specifies the Hot Standby Router Protocol (HSRP). The goal of the protocol is to allow hosts to appear to use a single router and to maintain connectivity even if the actual first hop router they are using fails. Multiple routers participate in this protocol and in concert create the illusion of a single virtual router. The protocol insures that one and only one of the routers is forwarding packets on behalf of the virtual router. End hosts forward their packets to the virtual router.

The router forwarding packets is known as the active router. A standby router is selected to replace the active router should it fail. The protocol provides a mechanism for determining active and standby routers, using the IP addresses on the participating routers. If an active router fails a standby router can take over without a major interruption in the host's connectivity. This memo also discusses the ARP, MAC address, and security issues with this protocol.

TABLE OF CONTENTS

1. Introduction

The Hot Standby Router Protocol, HSRP, provides a mechanism which is designed to support non-disruptive failover of IP traffic in certain circumstances. In particular, the protocol protects against the failure of the first hop router when the source host cannot learn the IP address of the first hop router dynamically. The protocol is designed for use over multi-access, multicast or broadcast capable LANs (e.g., Ethernet). HSRP is not intended as a replacement for existing dynamic router discovery mechanisms and those protocols should be used instead whenever possible [1]. A large class of legacy host implementations that do not support dynamic discovery are capable of configuring a default router. HSRP provides failover services to those hosts.

All of the routers participating in HSRP are assumed to be running appropriate IP routing protocols and have a consistent set of routes. The discussion of which protocols are appropriate and whether routing is consistent in any given situation is beyond the scope of this specification.

Using HSRP, a set of routers work in concert to present the illusion of a single virtual router to the hosts on the LAN. This set is known as an HSRP group or a standby group. A single router elected from the group is responsible for forwarding the packets that hosts send to the virtual router. This router is known as the active router. Another router is elected as the standby router. In the event that the active router fails, the standby assumes the packet forwarding duties of the active router. Although an arbitrary number of routers may run HSRP, only the active router forwards the packets sent to the virtual router.

To minimize network traffic, only the active and the standby routers send periodic HSRP messages once the protocol has completed the election process. If the active router fails, the standby router takes over as the active router. If the standby router fails or becomes the active router, another router is elected as the standby router.

On a particular LAN, multiple hot standby groups may coexist and overlap. Each standby group emulates a single virtual router. For each standby group, a single well-known MAC address is allocated to the group, as well as an IP address. The IP address SHOULD belong to the primary subnet in use on the LAN, but MUST differ from the addresses allocated as interface addresses on all routers and hosts on the LAN, including virtual IP addresses assigned to other HSRP groups.

If multiple groups are used on a single LAN, load splitting can be achieved by distributing hosts among different standby groups.

The remainder of this specification discusses the operation of a single standby group. In the case of multiple groups, each group operates independently of other groups on the LAN and according to this specification. Note that individual routers may participate in multiple groups. In this case, the router maintains separate state and timers for each group.

2 Conditions of Use

US Patent number 5,473,599 [2], assigned to Cisco Systems, Inc. may be applicable to HSRP. If an implementation requires the use of any claims of patent no. 5,473,599, Cisco will license such claims on reasonable, nondiscriminatory terms for use in practicing the standard. More specifically, such license will be available for a one-time, paid up fee.

3 Scope

This document describes the packets, messages, states, and events used to implement the protocol. It does not discuss network management or internal implementation issues.

3.1 Terminology

The language conventions of RFC 2119 [3] are used in this document.

4 Definitions

Active Router - the router that is currently forwarding packets for the virtual router

Standby Router - the primary backup router

Standby Group - the set of routers participating in HSRP that jointly emulate a virtual router

Hello Time - the interval between successive HSRP Hello messages from a given router

Hold Time - the interval between the receipt of a Hello message and the presumption that the sending router has failed

5 Protocol

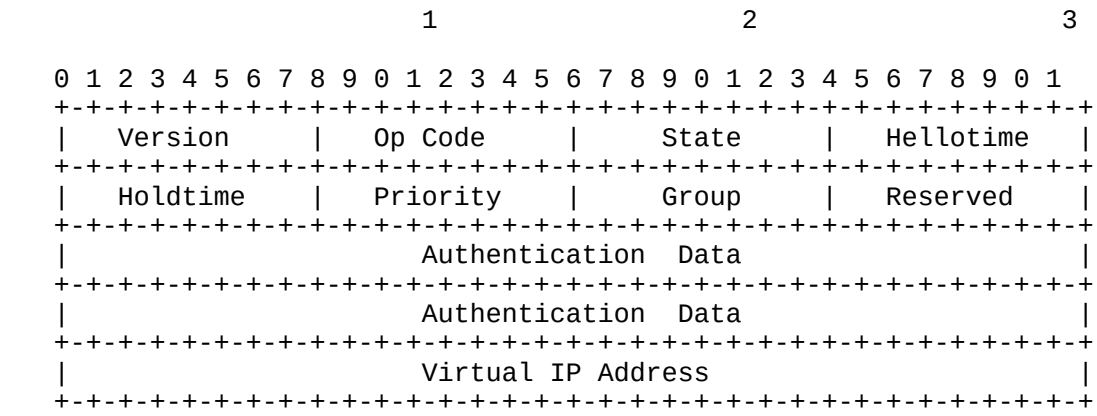
Within a standby group, the routers periodically advertise state information using various messages.

5.1 Packet formats

The standby protocol runs on top of UDP, and uses port number 1985. Packets are sent to multicast address 224.0.0.2 with TTL 1.

Routers use their actual IP address as the source address for protocol packets, not the virtual IP address. This is necessary so that the HSRP routers can identify each other.

The format of the data portion of the UDP datagram is:



Version: 1 octet

The version of the HSRP messages. This document describes version 0.

Op Code: 1 octet

The Op Code describes the type of message contained in this packet. Possible values are:

0 - Hello
1 - Coup
2 - Resign

Hello messages are sent to indicate that a router is running and is capable of becoming the active or standby router.

Coup messages are sent when a router wishes to become the active router.

Resign messages are sent when a router no longer wishes to be the active router.

State: 1 octet

Internally, each router in the standby group implements a state machine. The State field describes the current state of the router sending the message. Details on the individual states are described below. Possible values are:

```
  0 - Initial
  1 - Learn
  2 - Listen
  4 - Speak
  8 - Standby
 16 - Active
```

Hellotime: 1 octet

This field is only meaningful in Hello messages. It contains the approximate period between the Hello messages that the router sends. The time is given in seconds.

If the Hellotime is not configured on a router, then it MAY be learned from the Hello message from the active router. The Hellotime SHOULD only be learned if no Hellotime is configured and the Hello message is authenticated. A router that sends a Hello message MUST insert the Hellotime that it is using in the Hellotime field in the Hello message. If the Hellotime is not learned from a Hello message from the active router and it is not manually configured, a default value of 3 seconds is RECOMMENDED.

Holdtime: 1 octet

This field is only meaningful in Hello messages. It contains the amount of time that the current Hello message should be considered valid. The time is given in seconds.

If a router sends a Hello message, then receivers should consider that Hello message to be valid for one Holdtime. The Holdtime SHOULD be at least three times the value of the Hellotime and MUST be greater than the Hellotime. If the Holdtime is not configured on a router, then it MAY be learned from the Hello message from the active router. The Holdtime SHOULD only be learned if the Hello message is authenticated. A router that sends a Hello message MUST insert the Holdtime that it is using in the Holdtime field in the Hello message.

A router which is in active state MUST NOT learn new values for the Hellotime and the Holdtime from other routers, although it may continue to use values which it learned from the previous active router. It MAY also use the Hellotime and Holdtime values learned through manual configuration. The active router MUST NOT use one configured time and one learned time. If the Holdtime is not learned and it is not manually configured, a default value of 10 seconds is RECOMMENDED.

Priority: 1 octet

This field is used to elect the active and standby routers. When comparing priorities of two different routers, the router with the numerically higher priority wins. In the case of routers with equal priority the router with the higher IP address wins.

Group: 1 octet

This field identifies the standby group. For Token Ring, values between 0 and 2 inclusive are valid. For other media values between 0 and 255 inclusive are valid.

Authentication Data: 8 octets

This field contains a clear-text 8 character reused password.

If no authentication data is configured, the RECOMMENDED default value is 0x63 0x69 0x73 0x63 0x6F 0x00 0x00 0x00.

Virtual IP Address: 4 octets

The virtual IP address used by this group.

If the virtual IP address is not configured on a router, then it MAY be learned from the Hello message from the active router. An address SHOULD only be learned if no address was configured and the Hello message is authenticated.

5.2 Operational parameters

The following information MUST be known to each router in the standby group. The mechanisms used to determine this information are outside of the scope of this document.

Standby group number

Virtual MAC address

Priority

Authentication Data

Hellotime

Holdtime

The following information MUST be known to at least one router in each standby group and MAY be known by any of the other routers in the group.

Virtual IP Address

The following information MAY be configured on any router:

Preemption capability

If a router has higher priority than the active router and preemption is configured, it MAY take over as the active router using a Coup message.

5.3 States

Each router in the group participates in the protocol by implementing a simple state machine. This specification describes the externally visible behavior of this state machine. Implementations MAY vary their internal implementations within the functional description of the state machine.

All routers begin in the Initial state. This section discusses the intent of each state. For specific details on the actions taken in each state, please see the state transition table in section 5.7.

1. Initial

   This is the starting state and indicates that HSRP is not running. This state is entered via a configuration change or when an interface first comes up.

2. Learn

   The router has not determined the virtual IP address, and not yet seen an authenticated Hello message from the active router. In this state the router is still waiting to hear from the active router.

3. Listen

   The router knows the virtual IP address, but is neither the active router nor the standby router. It listens for Hello messages from those routers.

4. Speak

   The router sends periodic Hello messages and is actively participating in the election of the active and/or standby router. A router cannot enter Speak state unless it has the virtual IP address.

5. Standby

   The router is a candidate to become the next active router and sends periodic Hello messages. Excluding transient conditions, there MUST be at most one router in the group in Standby state.

6. Active

   The router is currently forwarding packets that are sent to the group's virtual MAC address. The router sends periodic Hello messages. Excluding transient conditions, there MUST be at most one router in Active state in the group.

5.4 Timers

Each router maintains three timers, an Active timer, a Standby timer, and a Hello timer.

The Active timer is used to monitor the active router. The active timer is started anytime an authenticated Hello message is seen from the active router. It is set to expire in the Holdtime seen in the Hello message.

The Standby timer is used to monitor the standby router The Standby timer is started anytime an authenticated Hello message is seen from the standby router. It is set to expire in the Holdtime seen in the Hello message.

The Hello timer expires once per Hellotime period. If the router is in Speak, Standby, or Active states, it should generate a Hello message upon Hello timer expiry. The Hello timer MUST be jittered.

5.5 Events

These are the events in the HSRP finite state machine.

a - HSRP is configured on an enabled interface.

b - HSRP is disabled on an interface or the interface is disabled.

c - Active timer expiry. The Active timer was set to the Holdtime when the last Hello message was seen from the active router.

d - Standby timer expiry. The Standby timer was set to the Holdtime when the last Hello message was seen from the standby router.

e - Hello timer expiry. The periodic timer for sending Hello messages has expired.

f - Receipt of a Hello message of higher priority from a router in Speak state.

g - Receipt of a Hello message of higher priority from the active router.

h - Receipt of a Hello message of lower priority from the active router.

i - Receipt of a Resign message from the active router.

j - Receipt of a Coup message from a higher priority router.

k - Receipt of a Hello message of higher priority from the standby router.

l - Receipt of a Hello message of lower priority from the standby router.

5.6 Actions

This section specifies the actions to be taken as part of the state machine.

A Start Active Timer
If this action occurred as the result of the receipt of a an authenticated Hello message from the active router, the Active timer is set to the Holdtime field in the Hello message. Otherwise the Active timer is set to the current Holdtime value in use by this router. The Active timer is then started.

B Start Standby Timer
If this action occurred as the result of the receipt of an authenticated Hello message from the standby router, the Standby timer is set to the Holdtime field in the Hello message. Otherwise the Standby timer is set to the current hold time value in use by this router. The Standby timer is then started.

C Stop Active Timer
The Active timer is stopped.

D Stop Standby Timer
The Standby timer is stopped.

E Learn Parameters
This action is taken when an authenticated message is received from the active router. If the virtual IP address for this group was not manually configured, the virtual IP address MAY be learned from the message. The router MAY learn Hellotime and Holdtime values from the message.

F Send Hello Message
The router sends a Hello message with its current State, Hellotime and Holdtime.

G Send Coup Message
The router sends a Coup message to inform the active router that there is a higher priority router available.

H Send Resign Message
The router sends a Resign message to allow another router to become the active router.

I Send Gratuitous ARP Message
The router broadcasts an ARP response packet advertising the group's virtual IP address and virtual MAC address. The packet is sent using the virtual MAC address as the source MAC address in the link layer header, as well as within the ARP packet.

5.7 State Transitions

This table describes the state transitions of the state machine. For each event and current state of the router, the router MUST perform the set of actions specified and transition to the designated state. If no action is specified, no action should be taken. If no state change is specified, no state change should be performed.

The notation used in this table has the specified set of actions listed as letters corresponding to the actions listed in section 5.6. The next state is listed as a number as specified in section 5.3. A slash ('/') separates the actions and states. Certain state transitions have alternatives which depend on external state. Alternatives are separated by a '|'. See the attached notes for details on these transitions.

States

Event	1 Initial	2 Learn	3 Listen	4 Speak	5 Standby	6 Active
a	AB/2|3+					
b		CD/1	CD/1	CD/1	CD/1	CDH/1
c			AB/4		CDFI/6	
d			B/4	D/5		
e				F	F	F
f				B/3	B/3	
g		EAB/3	EA	EA	EA	AB/4
h		EAB/3	A|BGFI/6*	A|BGFI/6*	A|BGFI/6*	G
i			AB/4	A	CFI/6	
j						ABH/4
k			B	B/3	B/3	B
l			B/4	D/5		B

Notes

\+ If the virtual IP address is configured, set state 3 (Listen) If the virtual IP address is not configured, set state 2 (Learn). In either case do actions A and B.

* If the router is configured to preempt do actions B, G, F, and I and set state to 6 (Active). If the router is not configured to preempt do actions A with no state change.

6 MAC Address Considerations

6.1 General

Each HSRP group has an associated well known virtual MAC address. On token ring networks, these addresses are actually functional addresses. The three addresses 0xC0 0x00 0x00 0x01 0x00 0x00, 0xC0 0x00 0x00 0x02 0x00 0x00, and 0xC0 0x00 0x00 0x04 0x00 0x00 correspond to groups 0, 1, and 2 respectively.

On other media, the virtual MAC addresses are 0x00 0x00 0x0C 0x07 0xAC XX where XX represents the HSRP group number. Routers which implement HSRP SHOULD use well-known HSRP MAC addresses as the group's virtual MAC address whenever possible.

The active router MUST accept and forward traffic that is destined for the group's virtual MAC address. It MUST stop accepting or forwarding such traffic when the router leaves the Active state.

If and only if the router is in the Active state, the router MUST use the group's virtual MAC address as the source MAC address for its Hello messages. This is necessary in order to allow learning bridges to be able to determine which LAN segment the virtual MAC address currently belongs to.

For each group, there is one virtual IP address and one virtual MAC address. This is a desirable situation, since the ARP table entries in the end stations do not need to change over time as the HSRP active router moves from one router to another.

Additionally, for HSRP to work in bridging environments, the bridges must be able to quickly update themselves as the virtual MAC address "moves". Although learning bridges typically are able to do this, some have been known to have problems with this. It is RECOMMENDED that only true learning bridges be used with HSRP.

The movement of the virtual MAC address can cause further undesirable side effects in environments where additional state is tied to the MAC address. For example on Token Ring, if Source Route Bridging is in use, a RIF will be stored with the virtual MAC address in a host's RIF cache. The RIF indicates the path and final ring used to reach the MAC address. As routers transition into Active state, they will not be able to affect the RIF caches on the hosts on the bridged ring. This may lead to packets being bridged to the ring for the previous active router.

In such circumstances, a router MAY use its normal MAC addresses as the virtual MAC address. This method of operation is strongly discouraged. In this mode, the virtual IP address will map to a different MAC address over time. This can create problems for end stations, since ARP tables assume a relatively static mapping between MAC address and IP address. These ARP tables are normally updated when the end stations receive the gratuitous ARP responses generated by a router that enters the active state.

6.2 Address Filter

As noted, routers currently emulating a virtual router adopt their group's MAC and IP addresses. MAC addresses are typically provided in an address filter or 'list' of MAC addresses in a router's interface controller. It is desirable for routers to be able to add one or more virtual MAC addresses to their controllers' MAC address filter while maintaining their primary MAC addresses.

Unfortunately, some interface controllers support address filtering for only one unicast MAC address. Or, in the case of Token Ring, the functional address which HSRP should use is already in use for some other protocol. In these cases, such routers can still implement HSRP, but the protocol must change the interface's primary MAC address when assuming or relinquishing control as the active router.

This is potentially problematic because some traffic may otherwise wish to use the router's primary MAC address. However, the problem MAY be mitigated by having the router send out gratuitous ARP packets regarding its non-HSRP IP addresses. Through this, other network entities using IP should update their ARP tables to reflect that the router is now using a group virtual MAC address rather than its primary MAC address.

Some protocols may not be able to run simultaneously with the standby protocol due to the interface primary MAC address change. For example, DECnet phase IV and HSRP will not be able to run at the same time on some equipment.

6.3 ICMP Redirect

While running HSRP, it is important to prevent the host from discovering the primary MAC addresses of the routers in its standby group. Thus, any protocol that informs a host of a router's primary address should be disabled. Thus, routers participating in HSRP on an interface MUST NOT send ICMP redirects on that interface.

6.4 Proxy ARP

Typically, hosts learn the HSRP virtual IP address through the configuration of their default router. These hosts then send packets for destinations outside of the LAN to the virtual IP address. In some environments, hosts may instead make use of proxy ARP in order to route off of the LAN. In this case, the hosts use the MAC address that is supplied in proxy ARP responses. HSRP functionality is maintained if the proxy ARP responses specify the HSRP virtual MAC address.

If an HSRP router is configured to support proxy ARP with HSRP, then the router MUST specify the HSRP virtual MAC address in any proxy ARP responses it generates. These proxy ARP responses MUST not be suppressed based upon HSRP state. Suppression based upon state could result in lack of any proxy ARP response being generated, since these proxy ARP responses may be suppressed due to other reasons, such as split-horizon rules.

7. Security Considerations

This protocol does not provide security. The authentication field found within the message is useful for preventing misconfiguration. The protocol is easily subverted by an active intruder on the LAN. This can result in a packet black hole and a denial-of-service attack. It is difficult to subvert the protocol from outside the LAN as most routers will not forward packets addressed to the all-routers multicast address (224.0.0.2).

8. References

[1] Deering, S., "ICMP Router Discovery Messages", RFC 1256, September 1991.

[2] United States Patent. Patent Number : 5,473,599. Standby Router Protocol. Date of Patent: Dec. 5, 1995.

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

9. Authors' Addresses

Tony Li
Juniper Networks, Inc.
3260 Jay St.
Santa Clara, CA 95054

Phone: (408) 327-1900
EMail: tli@juniper.net

Bruce Cole
Juniper Networks, Inc.
3260 Jay St.
Santa Clara, CA 95054

Phone: (408) 327-1900
EMail: cole@juniper.net

Phil Morton
Cisco Systems
170 Tasman Dr.
San Jose, CA 95143

Phone: (408) 526-7632
EMail: pmorton@cisco.com

Dawn Li
Cisco Systems
170 Tasman Dr.
San Jose, CA 95143

Phone: (408) 527-2014
EMail: dawnli@cisco.com

10. Full Copyright Statement