Slave Virtual Router Redundancy Protocol (SVRRP)

draft-zhang-vrrp-svrrp-00.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on April 19, 2010.

Copyright Notice

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents in effect on the date of publication of this document (http://trustee.ietf.org/license-info). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract
In this document, we propose a simplified VRRP protocol called the Slave Virtual Router Redundancy Protocol (SVRRP). The design objective of SVRRP is to specify an election protocol that dynamically assigns responsibility for a virtual router to one of the SVRRP routers on a LAN, which is exactly as same as that of VRRP. However, SVRRP executions do not exchange signaling packets and need the assistance of VRRP to perform their functionality appropriately. This approach can be used to improve the efficiency of VRRP routers in certain scenarios.

Conventions used in this document

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].

Table of Contents

1. Introduction.......................................................2
2. Terminology.....................................................................3
3. Motivating Scenario.......................................................3
4. SVRRP...........................................................................4
5. IANA Considerations.......................................................7
6. Acknowledgments.........................................................7
7. References.......................................................................7
8. Authors’ Addresses........................................................8

1. Introduction

In many typical scenarios (e.g., a VRRP router supporting multiple VLANs), a VRRP router needs to backup multiple virtual routes (VRs) simultaneously. In order to maintain the state of a VRRP instance (execution) for the election of a VR, the router needs to exchange VRRP signaling packets with other routers involved in the same selection. The bandwidth consumed by the router in transporting VRRP signaling packets is proportional to the number of the VRs which it supports. In order to improve the efficiency of the VRRP routers in such scenarios, a simplified VRRP, called Slave VRRP (SVRRP) is proposed. SVRRP need not exchange signaling messages. Actually, the state of a SVRRP instance is determined by another VRRP instance (which is referred to as a MVRRP instance in this case). Therefore, a VRRP router executing SVRRP needs in addition to execute MVRRP. This solution is based on the fact that the states of a set of VRRP instances located on a VRRP router are always identical and
transferred in a synchronized way, e.g., when they share a same physical interface.

2. Terminology


VRRP instance: a VRRP execution on a router for the backup of a virtual router.

MVRRP instance: a MVRRP execution on a router for the backup of a virtual router.

SVRRP instance: a SVRRP execution on a router for the backup of a virtual router.

VRRP backup group (VRRP BG): a collection of VRRP instances for the backup of a same virtual router.

MVRRP backup group (MVRRP BG): a collection of MVRRP instances for the backup of a same virtual router.

SVRRP backup group (SVRRP BG): a collection of SVRRP instances for the backup of a same virtual router.

Broadcast_Timer: a global timer that fire to trigger sending gratuitous ARP for SVRRP instances in Master.

Broadcast_interval: the interval of a Broadcast_Timer, default is 300 seconds.

3. Motivating Scenario

Figure 1 illustrates a motivating scenario where multiple hosts belonging to different VLANs connect to two VRRP routers (Router1 and Router2) through a switch. The packets from the hosts are routed to a same physical interface (e.g., Interface1) of the master VRRP router (e.g., Router1). Moreover, the physical interface is divided into multiple sub-interfaces; each is used to deal with the packets in a VLAN.
Because in principle a VR acts as a default router for hosts on a shared LAN, each of the routers should generate a VRRP instance for every VLAN respectively. As mentioned previously, a VRRP instance needs to maintain a state machine and periodically exchange signaling packets with others in the same VRRP back group. When the number of VRRP instances supported by a VRRP router is large, the resource (e.g., bandwidth, CPU, memory) occupied in transporting and processing signaling packets may influence the performance of the router negatively.

In this scenario, all the VRs share the same physical interface; the states of the VRRP instances located on a same VRRP router are always identical and change in a concurrent way. Therefore, it can be effective to select a VRRP BG as the MVRRP BG while the instances in other VRRP BGs are implemented with SVRRP. The state of a MVRRP instance is shared by the SVRRP instances on the same VRRP router, and any change in the state of the MVRRP instance can trigger identical changes in the states of the corresponding SVRRP instances. Therefore, Router1 and Router2 only need to exchange VRRP signaling packets for the MVRRP BG, irrespective of how many VRs they are supporting.

4. SVRRP

4.1. State Transition Diagram
4.2. State Descriptions

In the state descriptions below, the state names are identified by (state-name), and the packets are identified by all upper case characters.

There are two key events in the SVRRP state machine, the Startup event and the Shutdown event. A Startup event arises when the associated interface is available and the associate MVRRP instance is in the {Backup} or the {Master} state. A Shutdown event arises when the interface is unavailable or the associated MVRRP instance is in the {Initialize} state.

4.2.1. Initialize

A SVRRP instance in this state just waits for a Startup event. If a Startup event is received, then:

- If State of MVRRP is {Master}
  - If broadcast_timer is inactive, then:
    o set broadcast_timer to broadcast_interval.
    endif
  o Broadcast a gratuitous ARP request containing the virtual router MAC address for each IP address associated with the virtual router.
4.2.2. Backup

While in this state, MUST discard ADVERTISEMENT received.

- If receiving a Shutdown event:
  o Transition to the (Initialize) state
endif

- If the state of the associated MVRP instance is transferred into Master, then:
  o Broadcast a gratuitous ARP request containing the virtual router MAC address for each IP address associated with the virtual router.
  o Transition to the (Master) state
def

4.2.3. Master

While in this state, a SVRRP instance must discard the received ADVERTISEMENT packets and periodically send gratuitous ARP packets according to Broadcast_timer in order to update a cache of MAC address of switch and/or hosts.

- If a Shutdown event is received, then:
  o Transition to the (Initialize) state
endif
- If the state of the associated MVRRP instance is transferred into Backup, then:
  
  o Transition to the (Backup) state

endif

- If the Broadcast_timer fires, then:
  
  o Broadcast a gratuitous ARP request containing the virtual router MAC address for each IP address associated with the virtual router.

endif

5. IANA Considerations

No such considerations.

6. Acknowledgments

Thank Peter Smith for his kindly revision and precious comments.

7. References


Authors’ Addresses

Qiang Zhang
Huawei Technologies Co., Ltd
Huihong Building, No.91 Baixia Rd.,
Baixia District
Nanjing, 210001
P.R. China
Email: njzq@huawei.com

Dacheng Zhang
Huawei Technologies Co., Ltd
KuiKe Building, No.9 Xinxi Rd.,
Hai-Dian District
Beijing, 100085
P.R. China
Email: zhangdacheng@huawei.com