VRRP Protocol whitepaper

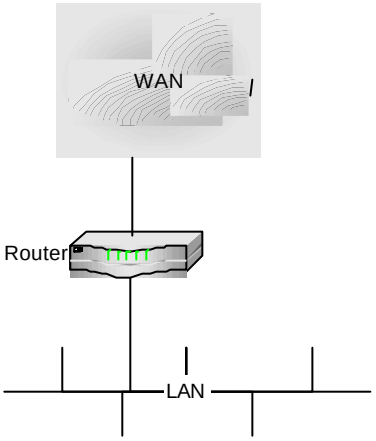
Contents

Introduction

Networks are the most important part of our communication system in every aspect of our lives. Without them we can't communicate to each other; can't get or send new information or can't withdraw our money from a bank subsidiary. Therefore network outages have to be as small as possible. There are networks which an outage will cause collapse of a business or loss of considerable amount of money. Such networks are willing or have to invest into redundancy solutions. There are many ways of minimizing network outages and every alternative has its pros and cons.

This document will describe a standardized VRRP protocol that deals with redundant gateways high availability system.

When a router is defined as a static default gateway and no other dynamic routing protocol or router discovery protocol is used, the gateway becomes a critical point in the network. If that router fails, that critical link would be broken and the LAN would be disconnected from other networks.



Pic.1: standard network

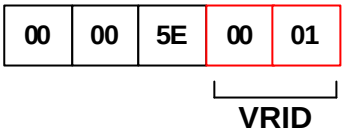
It, therefore, may be appropriate to set up other routers as backups that can serve as the static default gateway if necessary. But the problem is how to tell a client to use different gateway.

It is possible to do so by adding extra interfaces and routes on the client (which adds to complexity of the network and also slows down network communication at the client. Another possibility is to have a virtual gateway or router that will be independent on any physical device. This **Virtual Router** acts as a standalone network gateway for all clients on the LAN network.

The Virtual Router Redundancy Protocol (VRRP) eliminates the single point of failure inherent in the static default routed environment. VRRP specifies an election protocol that dynamically assigns responsibility for a virtual router to one of the network devices on a LAN. The Physical router controlling the IP address(es) associated with a virtual router is called the **Master**, and forwards packets sent to those IP addresses. When the Master becomes unavailable, a **backup** physical router takes the place of the Master. VRRP provides a function similar to a Cisco Systems, Inc. proprietary protocol named Hot Standby Router Protocol (HSRP) and to a Digital Equipment Corporation, Inc. proprietary protocol named IP Standby Protocol.

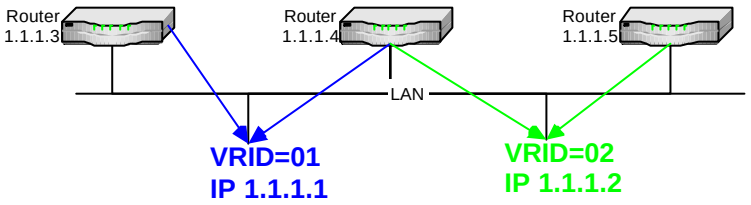
VRRP Protocol

VRRP specifies an election protocol to provide the virtual router function described earlier. All protocol messaging is performed using IP multicast datagrams. Each VRRP virtual router has a single well-known MAC address allocated to it. The virtual router MAC address is used as the source in all periodic VRRP messages sent by the Master router to enable bridge learning in an extended LAN.
The virtual router MAC address associated with a virtual router is an IEEE 802 MAC Address in the following format:



00-00-5E-00-01-{VRID} (in hex in internet standard bit-order)
The first three octets are derived from the IANA's OUI. The next two octets (00-01) indicate the address block assigned to the VRRP protocol. {VRID} is the VRRP Virtual Router Identifier. This mapping provides for up to 255 VRRP routers on a network.

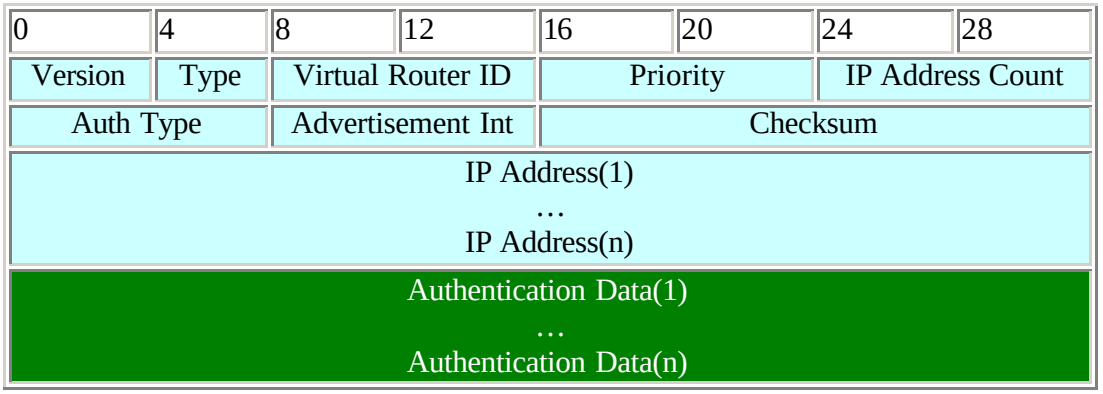
A virtual router is defined by its **virtual router identifier (VRID)** and a set of IP addresses. A VRRP router may associate a virtual router with its real addresses on an interface, and may also be configured with additional virtual router mappings and priority for virtual routers it is willing to backup. The mapping between VRID and addresses must be coordinated among all VRRP routers on a LAN.



To minimize network traffic, only the Master for each virtual router sends periodic VRRP Advertisement messages. A Backup router will not attempt to pre-empt the Master unless it has higher priority. This eliminates service disruption unless a more preferred path becomes available. It's also possible to administratively prohibit all preemption attempts. The only exception is that a VRRP router will always become Master of any virtual router associated with addresses it owns. If the Master becomes unavailable then the highest priority Backup will transition to Master after a short delay, providing a controlled transition of the virtual router responsibility with minimal service interruption.

VRRP defines three types of authentication providing simple deployment in insecure environments, added protection against misconfiguration, and strong sender authentication in security conscious environments.

VRRP packets are sent encapsulated in IP packets. They are sent to the IPv4 multicast address assigned to VRRP.



Picture 1: VRRP packet structure

IP header

Source Address
The primary IP address of the interface the packet is being sent from.
Destination Address
The IP multicast address as assigned by the IANA for VRRP is:
224.0.0.18
TTL
Time to Live (default is 255, other values should be discarded)
Protocol
The IP protocol number assigned by the IANA for VRRP is 112 (decimal).

VRRP Field Descriptions

Version

The version field specifies the VRRP protocol version of this packet.

Type

The type field specifies the type of this VRRP packet. The only packet type defined in this version of the protocol is:

1 ADVERTISEMENT

Virtual Rtr ID (VRID)

The Virtual Router Identifier (VRID) field identifies the virtual router this packet is reporting status for.

Priority

The priority field specifies the sending VRRP router's priority for the virtual router. Higher values equal higher priority. This field is an 8 bit unsigned integer field.

The priority value for the VRRP router that owns the IP address(es) associated with the virtual router MUST be 255 (decimal).

VRRP routers backing up a virtual router MUST use priority values between 1-254 (decimal). The default priority value for VRRP routers backing up a virtual router is 100 (decimal).

The priority value zero (0) has special meaning indicating that the current Master has stopped participating in VRRP. This is used to trigger Backup routers to quickly transition to Master without having to wait for the current Master to timeout.

IP Address Count

The number of IP addresses contained in this VRRP advertisement.

Authentication Type

The authentication type field identifies the authentication method being utilized. Authentication type is unique on a per interface basis. The authentication type field is an 8 bit unsigned integer. A packet with unknown authentication type or that does not match the locally configured authentication method MUST be discarded.

The authentication methods currently defined are:

0 - No Authentication
1 - Simple Text Password
2 - IP Authentication Header

No Authentication

The use of this authentication type means that VRRP protocol exchanges are not authenticated. The contents of the Authentication Data field should be set to zero on transmission and ignored on reception.

Simple Text Password

The use of this authentication type means that VRRP protocol exchanges are authenticated by a clear text password. The contents of the Authentication Data field should be set to the locally configured password on transmission. There is no default password. The receiver MUST check that the Authentication Data in the packet matches its configured authentication string. Packets that do not match MUST be discarded.

Note that there are security implications to using Simple Text password authentication, and one should see the Security Consideration section of this document.

IP Authentication Header

The use of this authentication type means the VRRP protocol exchanges are authenticated using the mechanisms defined by the IP Authentication Header [AUTH] using "The Use of HMAC-MD5-96 within ESP and AH" [HMAC]. Keys may be either configured manually or via a key distribution protocol.

If a packet is received that does not pass the authentication check due to a missing authentication header or incorrect message digest, then the packet MUST be discarded. The contents of the Authentication Data field should be set to zero on transmission and ignored on reception.

Advertisement Interval

The Advertisement interval indicates the time interval (in seconds) between ADVERTISEMENTS. The default is 1 second. This field is used for troubleshooting misconfigured routers.

Checksum

The checksum field is used to detect data corruption in the VRRP message.

The checksum is the 16-bit one's complement of the one's complement sum of the entire VRRP message starting with the version field. For computing the checksum, the checksum field is set to zero.

IP Address(es)

One or more IP addresses that are associated with the virtual router. The number of addresses included is specified in the "Count IP Addrs" field. These fields are used for troubleshooting misconfigured routers.

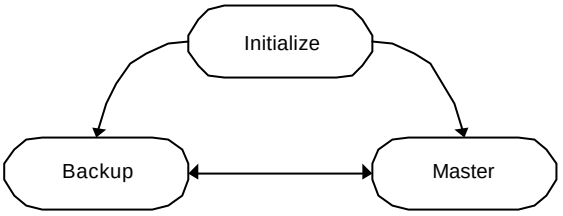
Authentication Data

The authentication string is currently only utilized for simple text authentication, similar to the simple text authentication found in the Open Shortest Path First routing protocol. It is up to 8 characters of plain text.

VRRP protocol states

There are 3 states defined for a VRRP protocol:

1. Initialize state (where device detects its state according to advertisement packets received)
2. Master state (state where device sends out advertisement packets)
3. Backup state (device listens to advertisements comparing its priority with priority advertised)



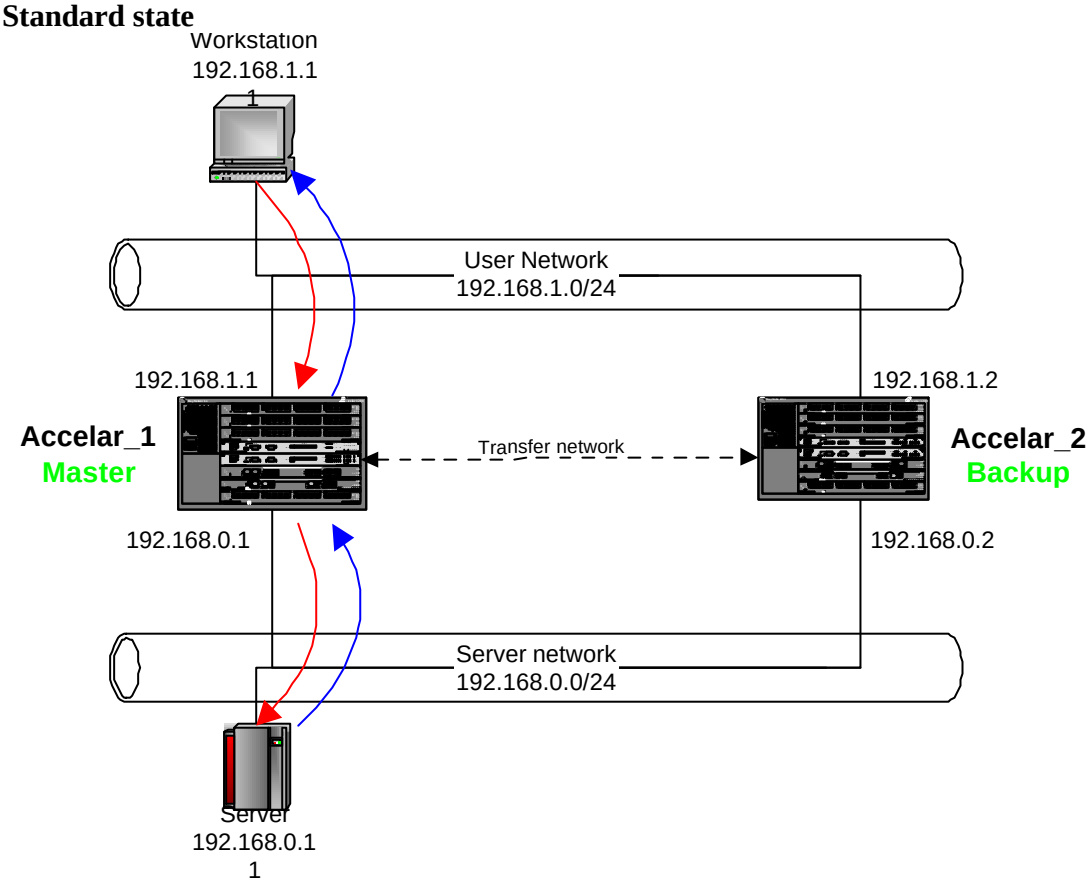
Specific definition of states and requirements for implementation of VRRP can be found in RFC 2338[13] part 6.2-6.4.

VRRP Implementations

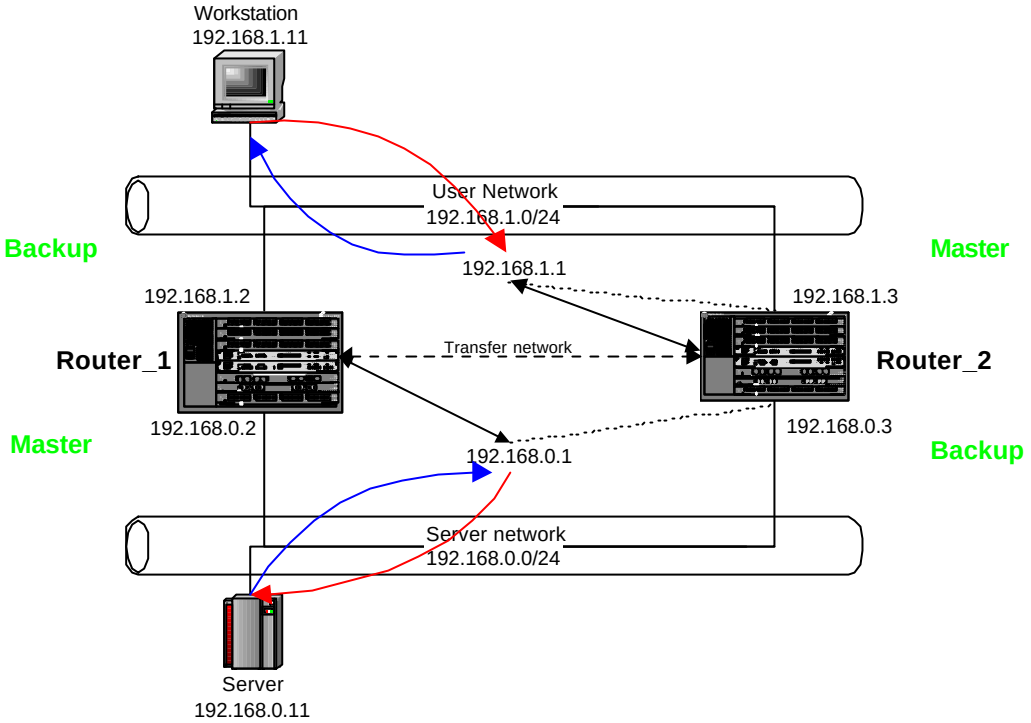
VRRP 2

This is the most common implementation seen in many network elements. Reasons for such VRRP deployment is not very clear to me, but it's weaknesses have to be considered in design of the network (in the case below it is the transfer network that's required).

Standard state



Failover state

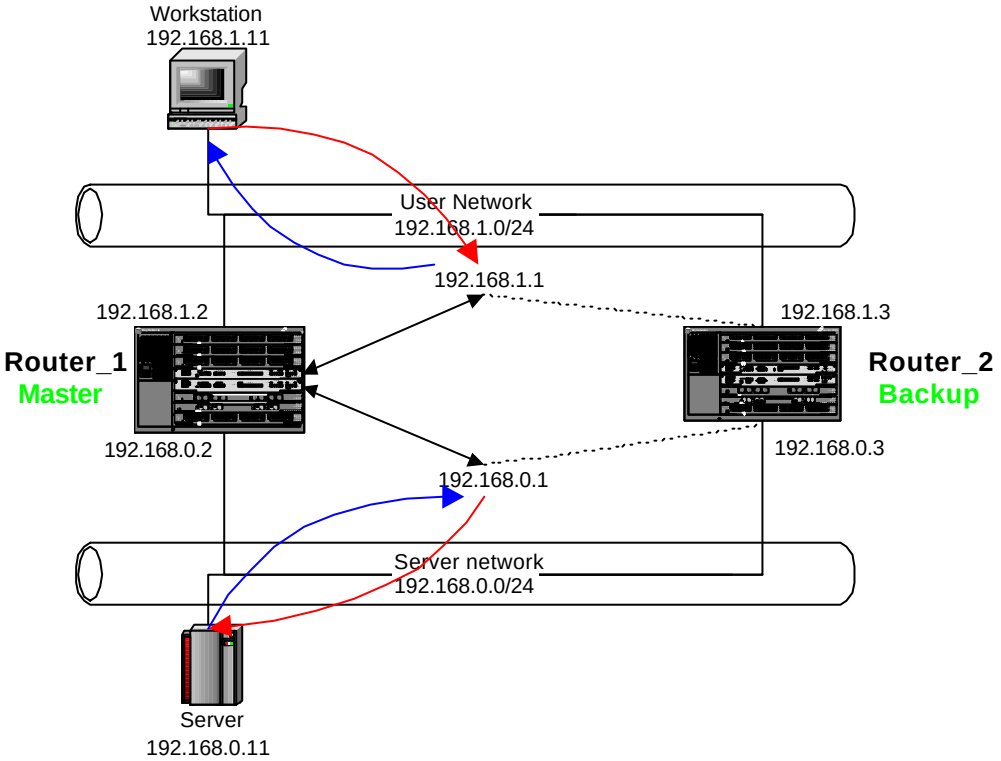


The major disadvantage of standard VRRP failover system is that it was designed for one side only. For router environment it requires an extra transfer network between routers to ensure that after failover the traffic will still reach its destination.

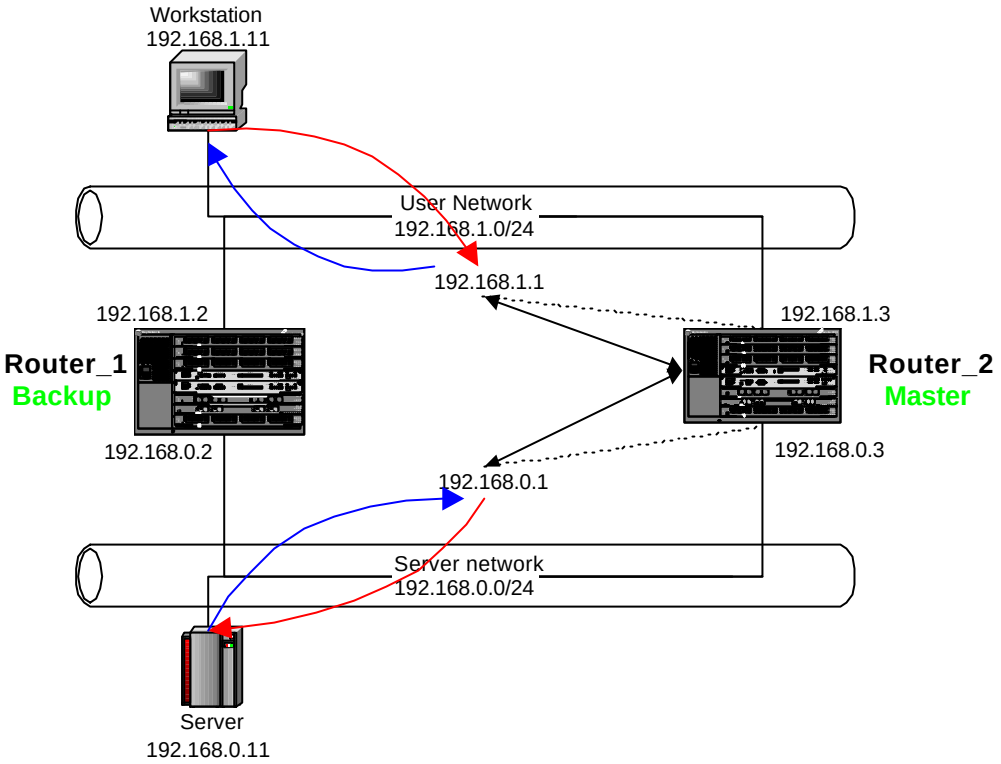
Monitored Circuit

This system is a more advanced VRRP, that not only monitors one network interface but also defined network flow. This way an outage of one user network interface on master router causes the network flow to be switched to backup router (also on server network interface master router switches to backup router). Standard VRRP system will keep user network gateway on backup router while on server network master router will remain the gateway and the traffic will flow only if there is a third interface directing traffic from master router to backup router and back.

Standard State



Failover state



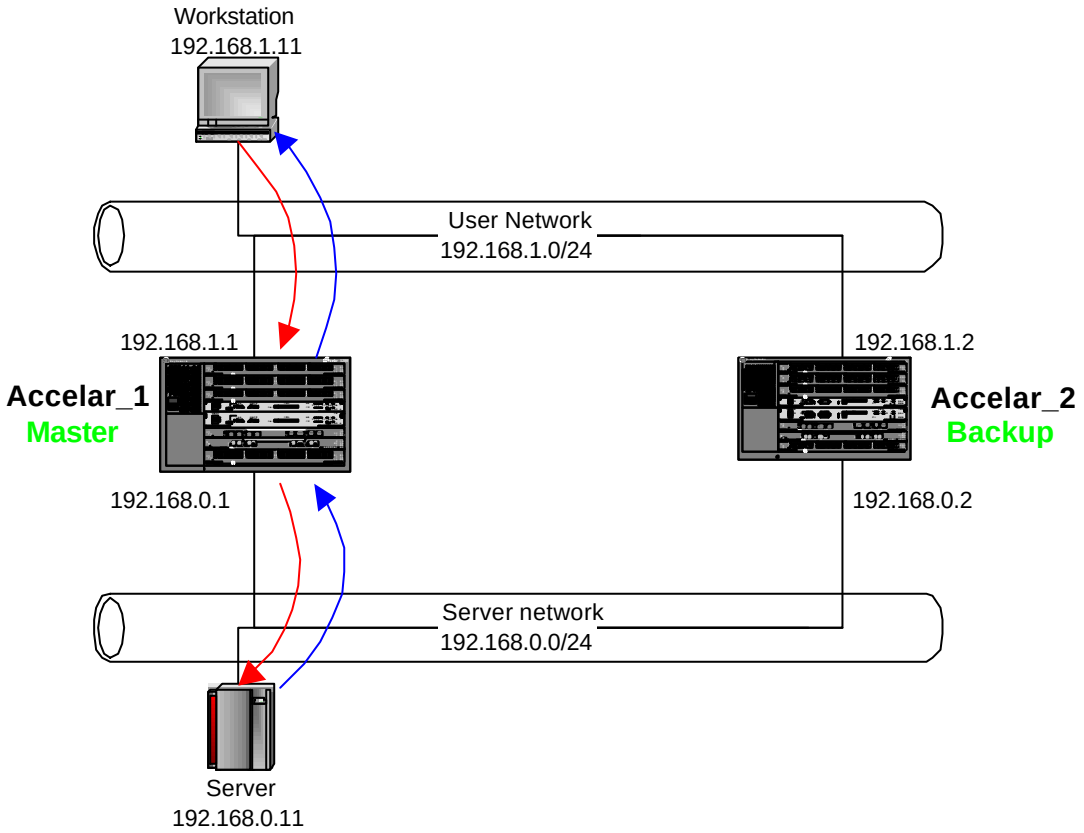
Here it is clear what advantages Monitored Circuit approach offers. Saving one interface on each router and extra utilization on each router is a clear PRO for this type of deployment. Unfortunately not all vendors use this VRRP approach (to my knowledge it is only NOKIA).

Deployment strategies

Implementation of VRRP system is not as easy as protocol looks like, as implementation has to be carefully planned and deployed. A simple error in the procedure may cause various problems that may show up much later after the implementation. Also all the steps during deployment should be carefully considered and executed, as priority exchange and failover is not always operational from the beginning and with some systems it may remain stuck in backup mode even if the priority is higher.

VRRP on routers

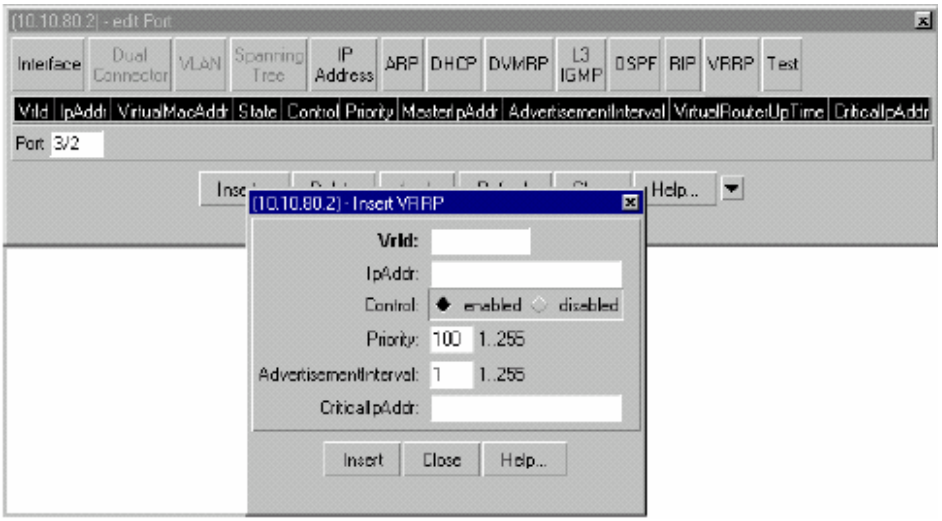
Router or gateway high availability is a necessity in important network areas where loss of connectivity is directly connected to loss of profit or business. VRRP system was primarily developed for router environment where routers are interconnecting several networks and each other. It is preferably used at network end-points, as intermediary systems prefer to use more flexible solution: dynamic routing.



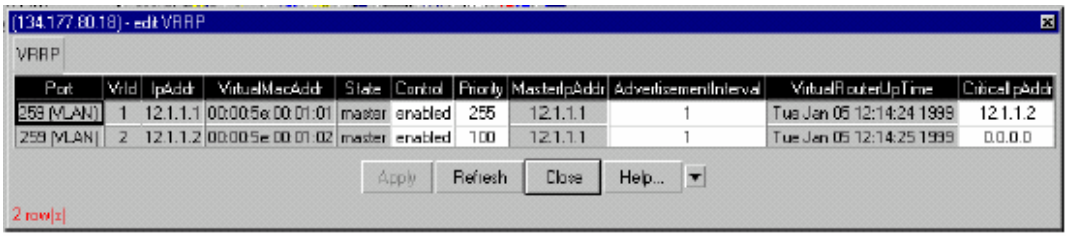
VRRP used by routers is usually a standard V2 version that doesn't support fail-over of other interfaces. In this example I'll use Nortel Accelar routers to show how to configure VRRP on a router:

To set up VRRP parameters:
• On a port, choose Edit > Port > VRRP.
• On a VLAN, choose VLAN > VLANs > Basic > IP > VRRP.
The Port VRRP window (below) and the VLAN VRRP window have the same fields described below:

- VrID :ID of the router
- IPAddr : IP address for the physical device or port
- Priority : priority of this router (higher is master)
- AdvertisementInterval : interval of sending advertisement packets
- CriticalAddr: IP address for the virtual router



To view all configured VRIDs for the switch, use the edit VRRP window. VRIDs cannot be added or deleted from this window.
From the main menu click on Routing > IP > VRRP. The window shown below opens



Configuration through CLI can also be used to reach the same result (without SNMP).
config ethernet [port] ip vrrp [VRID] followed by:

info	Displays the current port VRRP configuration
address `<ipaddr>`	Sets the IP address of the virtual router interface.
adver-int `<seconds>`	Sets the advertising interval (in seconds), the time interval between sending advertisement messages. The value must be the same on all participating routers. The range is 1 to 255, and the default is 1.
critical-ip `<ipaddr>`	Sets the critical IP address for VRRP. This address is an IP interface on the local router configured so that a change in its state causes a role switch in the virtual router (for example, from master to backup in case the interface went down).
delete	Deletes the VRRP from the port.
disable	Disables the VRRP on the port.
enable	Enables VRRP on the port.
priority `<prio>`	Sets the port VRRP priority (1 to 254) value to be used by this VRRP router. The default is 100. The value 255 is assigned to the router that owns the IP

	address associated with the virtual router.

Example:

```
Accelar-1200#config ethernet 3/3 ip/vrrp 2 info
                                Port 3/3 :
                             address : 192.168.1.1
                                 adver-int : 1
                             critical-ip : 0.0.0.0
                                 delete : N/A
                                 vrrp : enable
                                priority : 255
```

In order to see the configuration :

```
Accelar-1200#show ports info vrrp main
================================================================
Port Vrrp
================================================================
PORT_NUM VRRP_ID IP_ADDRESS   VIRTUAL_MAC_ADDR
----------------------------------------------------------------
3/3         2    191.168.1.1 00:00:5e:00:01:02

Accelar-1200# show ports info vrrp extended
================================================================================
Port Vrrp Extended
================================================================================
PORT STATE CONTROL PRIORITY MASTER_IPADDR ADVERTISE CRITICAL_IPADDR
--------------------------------------------------------------------------------
3/3 master enabled 255       192.168.1.1    1        0.0.0.0
```

config vlan [vlanid] ip vrrp [vrid] followed by:

info	Displays the current VLAN VRRP configuration
address <ipaddr>	Sets the IP address of the virtual router interface.
adver-int <seconds>	Sets the advertising interval (in seconds), the time interval between sending advertisement messages. The value must be the same on all participating routers. The range is 1 to 255, and the default is 1.
critical-ip <ipaddr>	Sets the critical IP address for VRRP. This address is an IP interface on the local router configured so that a change in its state causes a role switch in the virtual router (for example, from master to backup in case the interface went down).
delete	Deletes the VRRP from the VLAN.
disable	Disables the VRRP on the VLAN.
enable	Enables VRRP on the VLAN.
priority <prio>	Sets the VLAN VRRP priority (1 to 254) value to be used by this VRRP router. The default is 100. The value 255 is assigned to the router that owns the IP address associated with the virtual router.

Example:

```
Accelar-1200# config vlan 2 ip vrrp 1 info
                             address : 192.168.1.1
                                 adver-int : 1
                             critical-ip : 0.0.0.0
                                 delete : N/A
                              vrrp enable : enable
                                priority : 255
                                   set : N/A
                                 delete : N/A

Accelar-1200# show vlan info vrrp main
```

```
=================================================================
Vlan Vrrp
=================================================================
VLAN VRRP                VIRTUAL
ID    ID   IPADDR       MAC ADDR
-----------------------------------------------------------------
2     1  192.168.1.1 00:00:5e:00:01:01
```

```
Accelar-1200# show vlan info vrrp extended
=================================================================================
                                Vlan Vrrp Extended
=================================================================================
                                MASTER       ADVERTISE  CRITICAL
VID   STATE  CONTROL  PRIORITY IPADDR       INTERVAL   IPADDR
---------------------------------------------------------------------------------
2     master enabled  255      192.168.1.1  1         0.0.0.0
```

In order to debug VRRP configuration there are some special commands:

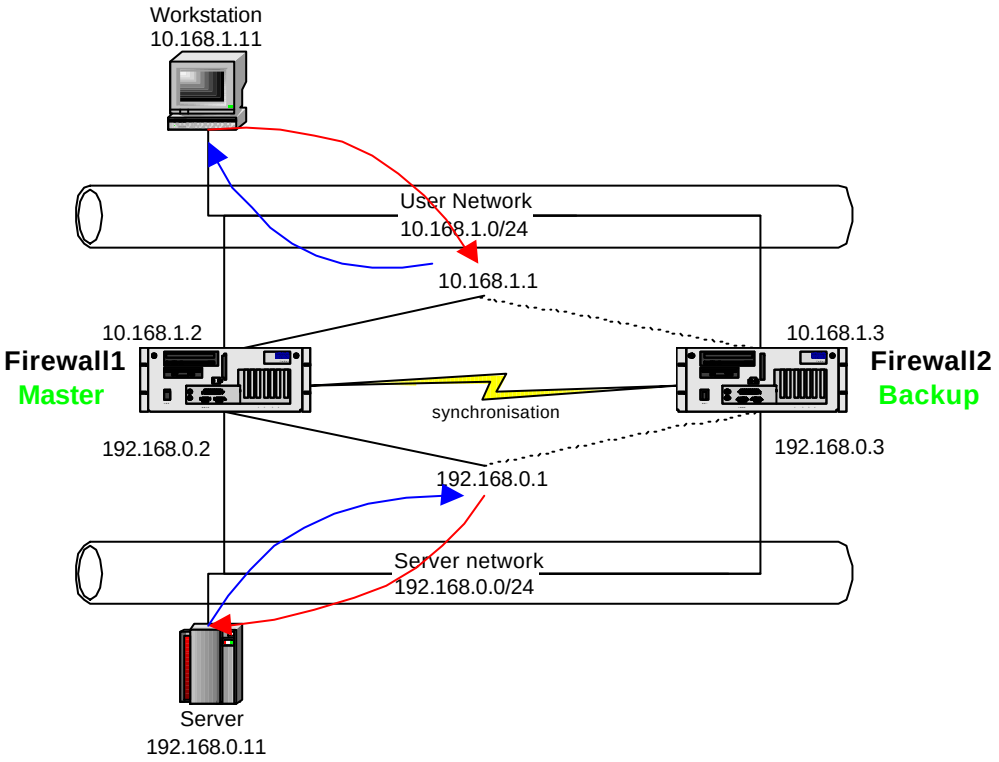
```
Accelar-1200# show ip vrrp info
=================================================================================
Vrrp Info
=================================================================================
VRID IP          MAC               STATE  CONTROL PRIO
---------------------------------------------------------------------------------
2    192.168.1.1 00:00:5e:00:01:02 Master Enabled 255
1    192.168.0.1 00:00:5e:00:01:01 Master Enabled 255
VRID MASTER      ADV UP                    CRITICAL
---------------------------------------------------------------------------------
2    192.168.1.1 1   0 day(s), 00:10:39   0.0.0.0
1    192.168.0.1 1   0 day(s), 00:11:08   0.0.0.0
```

```
Accelar-1200# show ip vrrp stats 1 100.100.100.1
=================================================================================
Vrrp Stats
=================================================================================
BECOME_MASTER ADVERITSE_RECEIVED CHECK_SUM_ERROR VERSION_ERROR
---------------------------------------------------------------------------------
0              0                   0               0
VRID_ERROR ADVERTISE_INT_ERROR TTL_ERROR PRIO_0_RECEIVED
---------------------------------------------------------------------------------
0          0                   0         0
PRIO_0_SENT INVALID_TYPE_ERROR ADDRESS_LIST_ERROR UNKNOWN_AUTHTYPE
---------------------------------------------------------------------------------
0           0                   0                 0
```

VRRP by firewalls

Firewall high availability is not as easy as router HA as there are other aspects requiring attention. If firewall is bound to an interface IP address it's not possible to use standard VRRP system without some modifications. Network address translation especially dynamic one is also not easy to implement. Monitored circuit system may allow us to forward packets from VRIP to our standard IP address and implementations of such firewalls is made possible. Checkpoint FW1 is not bound to interface IP address and can operate on both VRRP systems, as in case of interface change system modifies itself as well as it's filters.

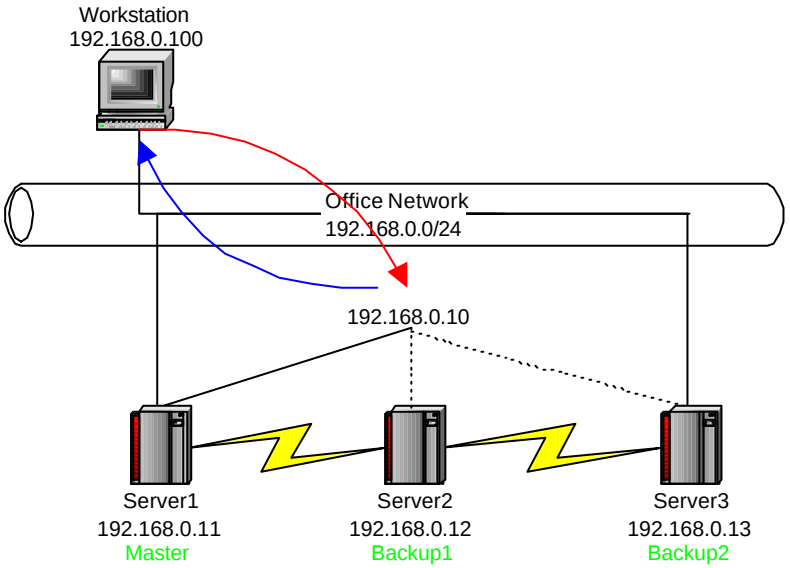
There is also another problem in case of fail-over: statefull packet filtering. Firewalls keep their IP flow states and if one firewall fails there is no possibility to recover connections in progress and workstation needs to reconnect TCP sessions to the server. Synchronization of packet state table can remove such problems, but preferably on extra interface without too much load on it.



VRRP on servers

High availability of important servers or applications is another possible place for VRRP to be used. There are already various system enhancements that allow VRRP to be used by operation systems [7].
One of such is VRRPd that can be downloaded at
http://w3.arobas.net/~jetienne/vrrpd/index.html
Or a bit newer modification on http://www.linuxvirtualserver.org/~acassen/

Implementation is done by installing a daemon that handles VRRP packets. These daemons are just ensuring high availability of the system and are not doing any load balancing or application synchronization. If there is a need to do load balancing there are other tools to do so as well (fx. Stonebeat [16], Radware [17], F5 Networks [18] or others). Load balancing can be also done by the application, but this is considered to be more an exception as applications are not designed for multihomed operations.
To utilize a fully redundant solution there needs to be a synchronization between servers or applications. Such synchronization causes a lot of network traffic and should be done on a special interface separated from other network traffic. When synchronizing application content to many servers multicasting may be more efficient saving network performance. There is no specific solution or protocol to do that as requirements from application or server point is too general.



VRRP test procedure

To fulfill HA requirements there needs to be a test procedure in place that ensures the system is deployed correctly. If the system works fine after deployment it doesn't mean that it will work fine after an hour or during fail-over or after fail-over state. Main part of the test procedure is the simulation of those states. To test the system we need a sniffer in place that may show us what is actually going on at every router interface. This may be accomplished by an expensive network tester (fx. NAI sniffer) or by a cheaper PC+unix+tcpdump alternative. Some routers already have such a tool built in (fx. Nokia).

Main state

System is fully functional and master is operational

1. ping the primary IP address (should be responding from the primary router)
2. packet flow should be going through primary router

Fail-over state

Master router is not working and backup router should take over the functionality

1. remove the connection on the primary router that is carrying the primary IP address
2. ping the primary IP address (should be responding within 3 seconds)
3. check the interface status on the backup router (backup router should contain 2 IP addresses and packets are flowing through it)

Return from fail-over state

Backup router is working and master router has recovered from failure.

1. Plug back the cable removed from the previous step
2. ping the primary IP address (should be responding within 3 seconds)
3. check the interface status on the backup router (the primary IP address should be gone) and on the primary router (packets should be using this interface)

Weak points

There is no protocol or implementation of it bulletproof, so that there is no undefined state that system may get into. Even VRRP protocol has minor issues that need to be addressed by protocol designers or implementators.

1. Fail-over problem on network layer (only one interface down)
2. Fail-over problem on application layer (application hanging)
3. Weak protocol authentication (VRIP takeover/DoS)
4. SNMP management security (section 4 of RFC 2787 [14])

Point 1.

Point 2.

Point 3. Weak protocol authentication (VRIP takeover/DoS)

Point 4.

Possibilities of improvement

Reference:

1. vrrp_nokia http://www.nokia.com.br/nic/pdf/vrrp.pdf
2. vrrp___ http://www.alliedtelesyn.co.nz/documentation/arrouter/231/pdf/vrrp.pdf
3. vrrp__ http://www.cisco.com/warp/public/471/vrrp.pdf
4. http://www.futsoft.com/pdf/VRRPfs001.pdf
5. http://lwn.net/2001/features/OLS/pdf/pdf/vrrpd.pdf
6. vrrp_ http://www.alliedtelesyn.co.nz/support/ar800/ar800-221/vrrp.pdf
7. http://keepalived.sourceforge.net/pdf/LVS-HA-using-VRRPv2.pdf
8. http://www.nortelnetworks.com/solutions/lan/collateral/ppvrrp.pdf
9. vrrp http://support.efficient.com/docs/pdf/vrrp.pdf
10. vrrp_pres http://csgrad.cs.vt.edu/~jxzhao/ECPE6504/presentations/VRRP.PDF
11. vrrpd homepage http://w3.arobas.net/~jetienne/vrrpd/
12. http://www.protocolsource.com/download/future_vrrp.PDF
13. http://www.ittc.ku.edu/~subhashiv/845/vrrp/presentation.ppt
14. RFC 2338, Virtual Router Redundancy Protocol
15. RFC 2787, Definitions of Managed Objects for the Virtual Router Redundancy Protocol
16. http://www.stonesoft.com/products/ServerCluster/
17. http://www.radware.com/content/products/index.asp
18. http://www.f5.com/f5products/
19.