CHAPTER 2

VRRP Overview

This chapter gives an overview of the VRRP as a redundancy protocol without going into the details of used messages, the details of the states the protocol specifies, and the details to be taken into account while operating the protocol. Chapter 4 provides these details. Here we first put VRRP in context to articulate the significance of the function VRRP protects; then we talk about the special circumstances that necessitate a protocol such as VRRP. After establishing the context, we introduce basic VRRP concepts by using a series of simple configurations for didactic purposes. This approach helps us to explore some basic characteristics of the protocol: What is the typical VRRP configuration for establishing some level of load sharing? How is M-to-N redundancy established with VRRP?

The coverage of the basic elements of the protocol makes it possible to study some more realistic cases from the last section of this chapter. The last section discusses some typical VRRP deployment configurations before summarizing the covered topics.

2.1 THE CASE FOR VRRP

To provide a usage context for VRRP we consider an enterprise network connecting a corporate office to multiple branch offices in different regions. Figure 2-1 represents such an enterprise network.

A *cloud*—that is, a network of unspecified type and topology, but typically a Wide Area Network (WAN)—connects the branch offices to the corporate office, the headquarters of the enterprise. This WAN may consist of leased circuits, an X.25 packet network, a Frame Relay, or an ATM network. Another possibility is a traditional IBM's Systems Networking Architecture (SNA). But the cloud may also be representing an IP-based network of networks: a private internet or Internet with a capital *I*.

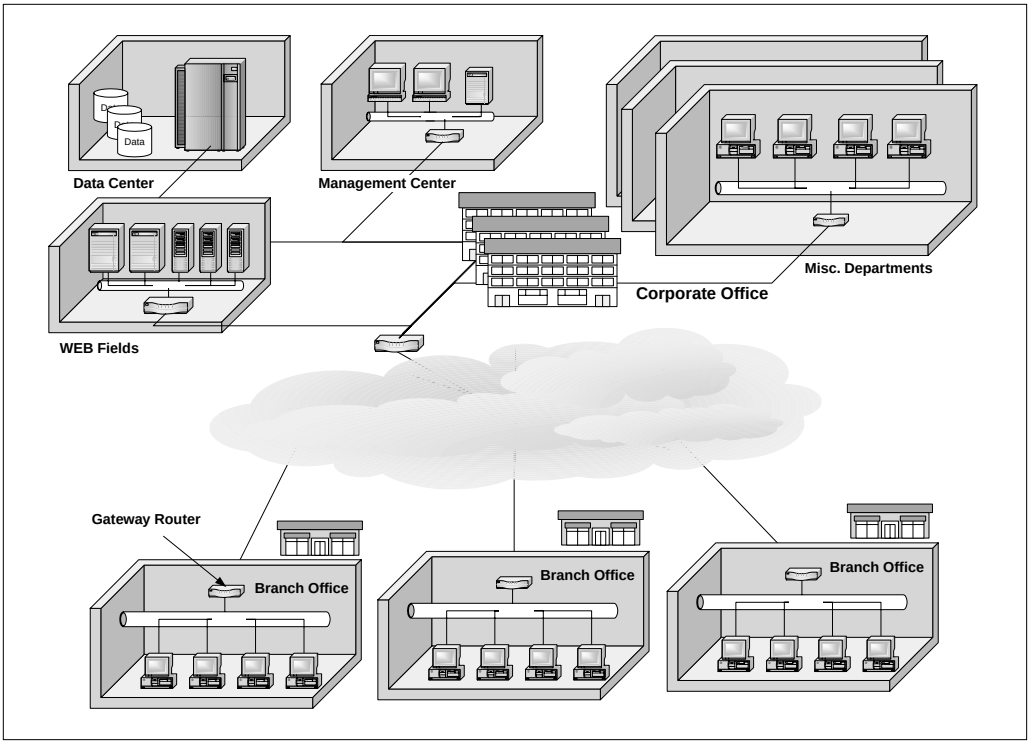

FIGURE 2-1. *An enterprise network*

The data center containing all business-critical databases resides in the corporate office, as does the network management center providing services essential to the operation of the corporate network. Finance department, payroll, human resources, and some other business functions are centralized to the headquarters. Different departments in the corporate office have their own interconnected LANs and are all connected to the external world through a device that we shall temporarily call a *gateway,* since it opens the gate to the external world. In our example, a router of some kind provides the gateway function. You may also hear the term *default router* used for this purpose. Another name for routers in this role is *first hop router,* since this is the first router the hosts need to use to reach any destination they cannot reach through direct routing.

Given the centralized setup of our illustration, branch offices of the enterprise depend heavily on the computer resources residing in the corporate office. Without access to these resources, the branch offices may fail to perform even their most basic functions.

The network settings and computer equipment of the branch offices are more modest. Depending on the specifics of the business, it may just contain a series of locally networked PCs. Connection to the cloud is established through a router acting as a gateway device, that is, a first hop router.

A Closer Look at the Branch Office

Now given the importance of the corporate resources that branch offices can only access through the cloud, the availability of the network services is extremely critical to the business. In our scenario, first hop routers constitute single points of failure.

VRRP, the protocol we are going to discuss in this book, provides a scheme intended to avoid this specific type of single points of failure. To understand how the need arises for this specific scheme and how VRRP basically achieves this avoidance, we zoom in to one of the branch offices.

In this branch office configuration, we have a series of general-purpose computers—PCs, workstations, laptops—interconnected via a LAN (typically Ethernet) to exchange local information, to share local resources. In this configuration, it is extremely clear that the first hop router constitutes the single point of failure for the network access of the branch office. In order to avoid this problem, the first step would be to introduce a redundant first hop router. Figure 2-2 depicts a configuration with two routers positioned to act in the first hop function.

Now several questions arise: How do these hosts decide which one of the routers (R1 or R2) to use initially or in normal circumstances? How do they switch to the available one should the one they were using fail? Hosts thus need a mechanism to decide which one of the hosts to use. They also need a mechanism to decide when to perform a switch. Moreover, when there is more than one backup router, they also need a mechanism to decide to which one of the backups to switch. Thus the questions at hand are threefold:

1. How do the hosts discover or select the first hop routers?
2. How do the hosts switch to backups when the master fails?
3. How do the hosts decide which backup to switch to when there is more than one?

There are various ways of implementing a mechanism to answer the above questions. The solution depends first of all on a series of assumptions surrounding the questions at hand. Maybe the most fundamental one of these assumptions is that the network depicted in Figure 2-2 (network N1) is an IP network, that the communication between the hosts and the router relies on IP, and that the router under discussion is an IP router. Note that the cloud (network N2) may be running a variety

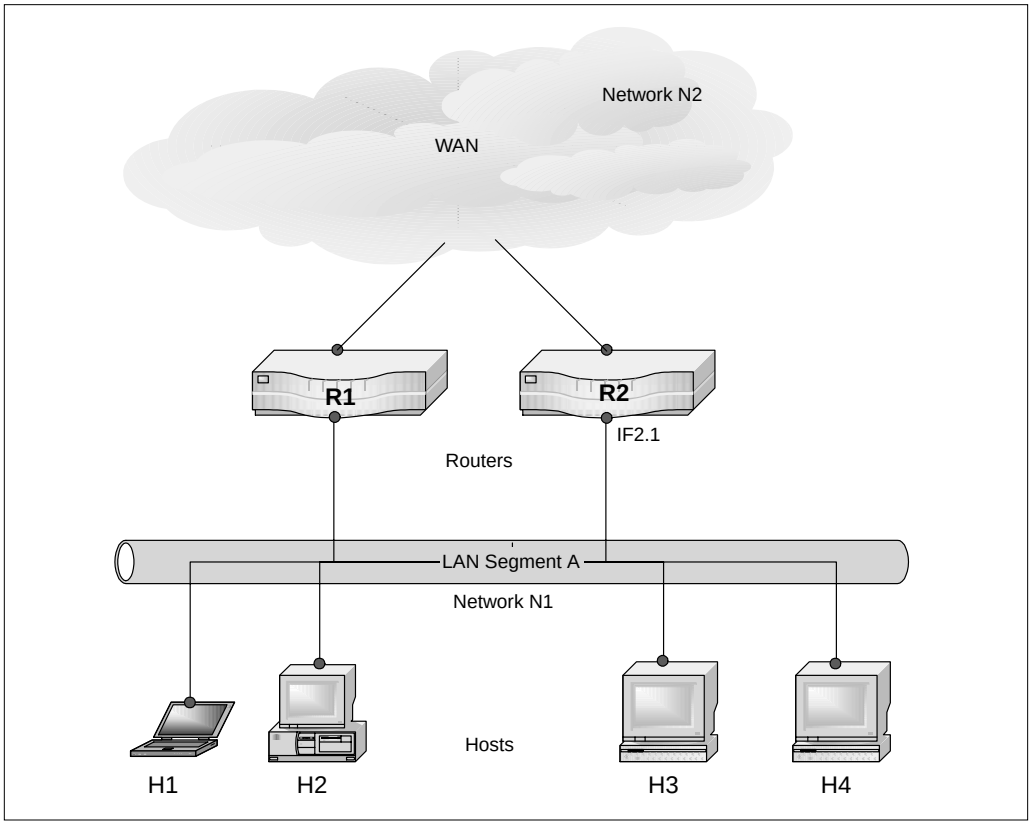

FIGURE 2-2. *Router pair*

of protocols: ATM, Frame Relay, and so on. Network N2 is most typically the network of a service provider, the transport mechanisms of which are transparent to the IP connectivity of the network N1.

Furthermore, there are additional questions that are of prime importance to network administrators in particular:

1. What is the status of backups when they act as backups? Are they just on standby, or are they also handling some portion of the traffic? What are the advantages and disadvantages of both approaches?
2. If the backups are just on standby, how is this wasteful use of resources justified?
3. If the backups are incorporated in a load-sharing arrangement, what are the mechanisms for achieving that?

Dynamic Routing

One category of solution consists of installing additional networking software within the hosts. Introducing dynamic routing software (such as RIP or OSPF) would come into this category. With the routing software enabled, hosts can behave as routers in addition to their end-system behavior. They can discover the first hop routers available in the LAN, for example, by listening to RIP updates on port 520. In case of failure, they can switch to alternate routers based on the routing protocol at hand. To follow the RIP example, they may receive a triggered event implying the failure of the current first hop router. Moreover, the metric mechanisms of the protocol can establish the switching priority among alternative first hop routers. By assigning different costs to different links, the network administrator can create a hierarchy of alternative first hop RIP routers.

This solution may be not the best or may not be feasible for several reasons. The hosts may not have the capacity to run routing software—in particular, a sophisticated and therefore demanding software such as OSPF. But even RIP in its passive mode, where the host can just listen to the updates without sending messages of its own, can be too much for an underpowered PC. Even if they were able to run RIP, RIP might turn out to be too slow to adapt to topology changes. In some cases, the implementation of these protocols may not be available at all on certain platforms.

Administrative concerns constitute another major obstacle to the use of dynamic routing protocols in the setting under discussion: the installment, the configuration, and the overall management of these protocols may be totally undoable given the centralization of the network management in our illustrative enterprise. Many administrators consider running routing protocols at the desktops a *management nightmare*. Keeping track of the routing software on different machines, making sure that they are all configured properly, they are all interoperable. . . . These are next to insurmountable challenges. Centralizing the routing intelligence at the edge of outbound pipe brings substantial simplicity to overall availability solutions in a way totally transparent to the end stations.

Security considerations may always be show stoppers in most networking issues, including running a dynamic protocol in a remote branch office ill equipped to handle basic protection requirements.

ICMP Discovery

As an alternative to dynamic routing, we can look at ICMP and consider running an ICMP router discovery client on our hosts. As a matter of fact, in recent times, some of the newer IP hosts use Router Discovery Protocol (RDP) to find a new router when a route becomes unavailable. A host that runs RDP listens for hello multicast messages from its configured router and uses an alternative router when it no longer receives those hello messages. The default timer values of RDP mean that it's not

suitable for quick detection of failure of the first hop, since the default advertisement rate is once every 7 to 10 minutes and the default lifetime is 30 minutes.

Moreover, this approach would require the active participations of all hosts. The increasing number of the hosts would require larger timer values to minimize the protocol overhead in the network. The larger timers, on the other hand, would lead to longer delays in the discovery of the failing neighbors, most important of the neighbors in the first hop router role. The result of these delays can be unacceptably long black hole periods.

Proxy Address Resolution Protocol (ARP)

Some IP hosts use proxy ARP to select a router. When a host runs proxy ARP, it sends an ARP request for the IP address of the remote host it wants to contact. A router, let us say R1 on the network, replies on behalf of the remote host and provides its own MAC address. With proxy ARP, the host behaves as if the remote host were connected to the same segment of the network. If the router R1 fails, the host continues to send packets destined for the remote host to the MAC address of R1, even though those packets have nowhere to go and are lost. You can either wait for ARP to acquire the MAC address of another router—say, R2 on the local segment—by sending another ARP request, or reboot the host to force it to send an ARP request. In either case, for a significant period of time the host can't communicate with the remote host, even though the routing protocol has converged, and R2 is prepared to transfer packets that would otherwise go through R1.

Dynamic Host Configuration Protocol (DHCP)

An interesting alternative is DHCP, a very popular and common method for providing configuration information to hosts on IP networks. A host running a DHCP client requests configuration information from a DHCP server when it boots onto the network. This configuration information typically comprises an IP address for the host and an IP address for a first hop router. Once configured, there is no mechanism within DHCP for switching to an alternative router if the default router fails. From the point of view of our stated problem, DHCP helps with discovery or selection of the first hop routers, but not with creating a redundancy scheme and establishing a switchover mechanism.

Static Configuration and VRRP

The final alternative under consideration would be reliance on the static routing. In this scenario, a network administrator would configure the IP address of the host as well as a first hop router as the default router for the host or use DHCP as discussed to obtain an IP address for the client and to learn the IP address of the default router.

This is by far the most feasible of the discussed alternatives. Static configurations are supported almost without exception by all TCP/IP implementations. This solution becomes even more attractive with the continuous deployment of DHCP clients and servers that substantially facilitates the configuration of the default routers.

The obvious shortcoming of this approach, though, is the same as the one we have discussed in our conclusion about DHCP. The static configuration helps with the discovery of the first hop router but does not help with the switchover or the selection of a master from the multiple backups. This very shortcoming leads to the creation of VRRP.

Having the first hop router configured as the default router establishes access to the external network, but another mechanism is needed to keep the access available. In the case of a failing default router, the hosts have no means for switching to an alternative router available in the network. Under these circumstances, unfortunately, the statically configured first hop router becomes the single point of failure for the network availability. VRRP enables additional routers to take over the role of a failing first hop router, thus helping them to avoid becoming the single point of failure for network services. Table 2-1 summarizes the advantages and disadvantages of the different approaches to protect first hop (gateway) routers.

TABLE 2-1. *Different Approaches to Protect the Gateway Routers*

SOLUTION	ADVANTAGES	DISADVANTAGES
Dynamic Routing	flexible: different failover hierarchies can be established	requires routing software in the hosts
	more failure points protected	demanding on the hosts
		challenging to administer
		security concerns
ICMP Discovery	may come with new IP stacks in the hosts	requires discovery software in the hosts with older IP stacks
		not responsive, may lead to black holes
Proxy ARP	no special software required in the host	not repsonsive, may lead to black holes
		may require rebooting of the hosts
DHCP		does not help with failover
VRRP with Static Routing	static route configuration commonly available with TCP/IP	protects only the default router's local interface but mechanism extendable to trigger a switchover in the case of other interfaces

(*continued*)

TABLE 2-1. *Different Approaches to Protect the Gateway Routers (continued)*

SOLUTION	ADVANTAGES	DISADVANTAGES
	does not require any special software in the host	applicable only to IP networks, but other protocols such as IPX on a VRRP interface able to be piggy-backed with IP for VRRP
	based on an industry standard	
	responsive or can be fine-tuned to be responsive	

2.2 BASIC CONCEPTS AND CONFIGURATIONS

Now that we have a context for the relevance of VRRP as well as the specific problem it is designed to solve, let us look at some basic elements VRRP uses to create the required redundancy or switchover mechanism. To do that, we revisit the LAN of the branch office discussed in Figure 2-2. On this LAN N1 network resides a series of hosts—H1, H2, H3, and H4—and two routers—R1 and R2—both positioned as potential first hop routers. As depicted in Figure 2-3, we assign R1 as the default router to all the hosts on N1 network.

This configuration requires that any message originating from H1, H2, H3, and H4 addressed to a destination outside N1 network be sent to R1. VRRP specifies a mechanism using which R2 starts acting as the default router when R1 fails, so that the hosts on N1 network do not become isolated. This section introduces basic concepts required for the high-level understanding of this switchover mechanism.

To establish the VRRP switchover mechanism, first of all, we need to run VRRP in R1 and R2. We refer to the routers running VRRP as *VRRP routers*. Using the mechanism specified by VRRP, we designate the VRRP routers R1 and R2 as the members of a VRRP *virtual router*. A VRRP virtual router consists of a group of VRRP routers that collaborate with each other to reduce the risk of having a single point of failure for a network service, in this specific case the function of a default first hop routing. Figure 2-4 depicts the simplest configuration for a virtual router. To refer to a virtual router, to be able to express the membership of VRRP routers to a specific virtual router, we need a label for identifying the group. VRRP calls this label *Virtual Router Identifier* or VRID. In our drawing, we use V1, V2, . . . Vn for VRIDs. For the sake of illustration, we use V1 to refer to the virtual router in Figure 2-4. In our specific example, we assign the role of master to R1 and the role of backup to R2 in the V1 virtual router using the mechanisms specified by VRRP.

Another point to emphasize is the object of protection. Since the IP addresses identify the connection of routers, of layer 3 elements to the network, we emphasize that VRRP protects the interfaces of a router providing default first hop services.

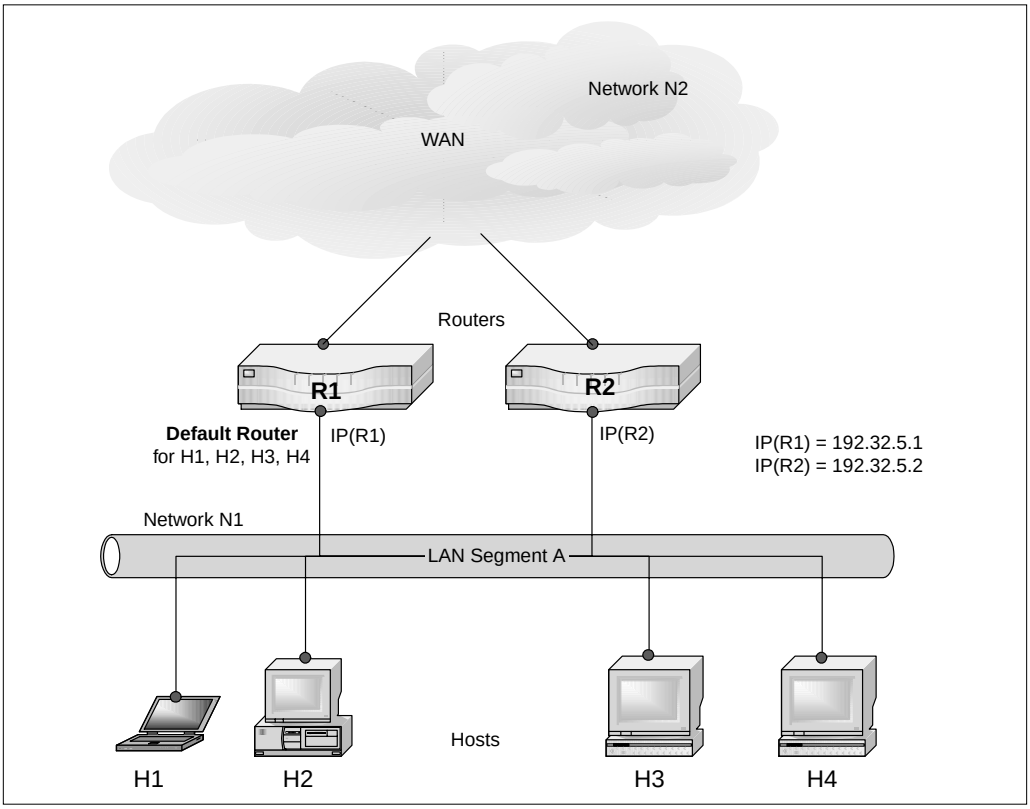

FIGURE 2-3. *Single default router*

Figure 2-4 graphically depicts all the considerations described so far. The oval shape in the drawing labeled with VRID V1 (or 37 as an illustration that VRID is an integer) indicates that R1 and R2 are members of the VRRP virtual router V1. By overlaying the oval shape on the lines originating from interfaces, we highlight that the objects of protection are the interfaces rather than routers. By putting the legend *master(V1)* under R1 and *backup(V1)* under R2, we indicate that R1 is the master in virtual router V1 and R2 is the backup.

According to this configuration, as long as the master, R1, is functional, all traffic destined to the external network gets directed to R1. But as soon as R1 fails, R2 takes over as the master and starts handling packets forwarded to the interface associated with IP(R1).

Figure 2-4 represents this situation. In this configuration, if R1 fails, R2 takes over the master responsibility and all external traffic gets directed now to R2.

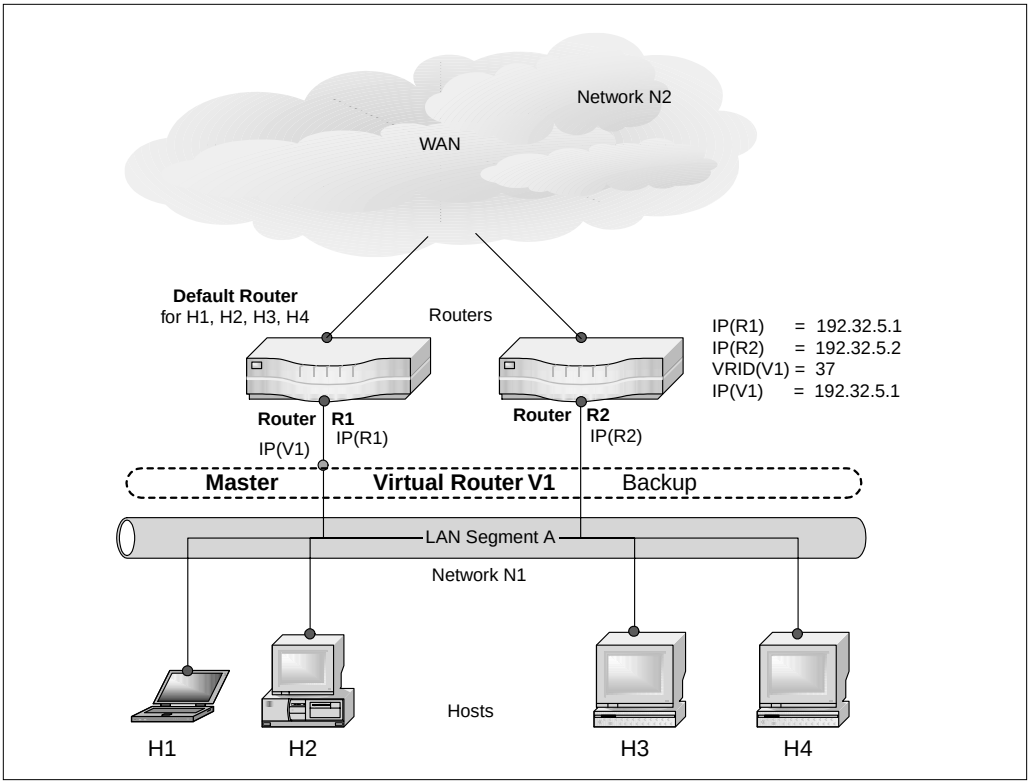

FIGURE 2-4. *One-sided protection*

2.3 LOAD SHARING

Note that in Figure 2-4, R2 router is completely idle during its backup periods. Its whole purpose in the network consists of being a backup for R1. R2 is a purely redundant device. In certain circumstances this may not be the best arrangement. Leaving a network device idle, in particular in the case of a robust master, may be considered underutilization or waste of a valuable resource. In such circumstances the network deployer may decide to assign R2 as a default router to some hosts on the LAN—say, H3 and H4. With this configuration traffic from H1, H2 is forwarded to R1 and traffic from H3 and H4 to R2.

The obvious advantage of this configuration is the establishment of a load-sharing scheme. With this configuration, the traffic originating from N1 network is not sent exclusively to one of the routers but is shared between R1 and R2. The traffic coming from H1 and H2 is handled by R1, and H3-H4 traffic is forwarded to R2.

But it must be realized that the virtual router configuration we have been discussing so far would not be of help to protect R2 in its default first hop router role, since being a member of a virtual router does not imply protection for the interfaces of a VRRP router. The protection needs to be explicitly set up.

To create this setup, we need to define two virtual routers, V1 and V2, and we need to define the opposite roles to our routers in V1 and V2. Figure 2-5 illustrates this setup.

In this setup R1 is defined as the master of the V1 and R2 as the backup. In V2, R2 is the master and R1 acts as the backup. With this configuration, we establish a load-sharing arrangement between R1 and R2; moreover, we create a mutual protection setup by having two routers acting as backups for each other. Note, however, that in the case of the failure of one of the routers, one can observe some degradation in the network service, unless the network is engineered accordingly.

Although many network deployers expressed to us their liking for the load-sharing aspect of this configuration and were agreeable to suffering some level of loss

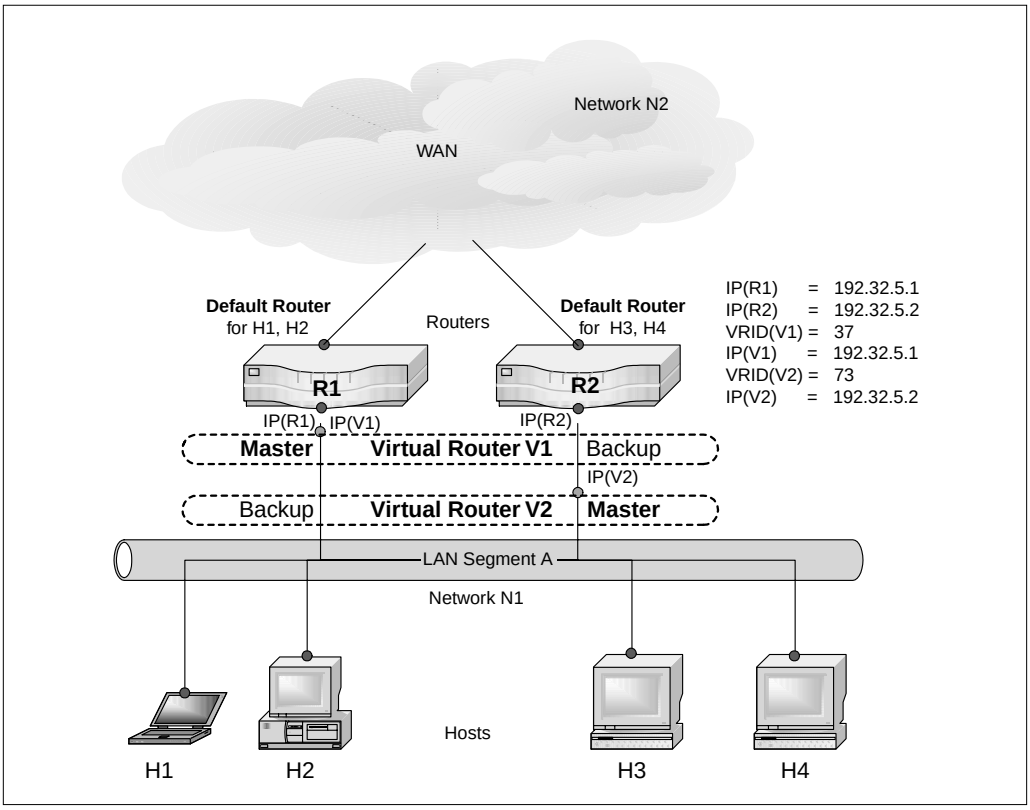

FIGURE 2-5. *Load sharing between VRRP routers*

in the quality of service, the service degradation was not acceptable to some other customers, and these others had rather a strong preference for a one-way protection scheme, even at the expense of leaving a resource completely idle during the backup periods. Our experience votes for load sharing. Often, backup equipment does not work when needed just because it is not kept in a fully working state. Any equipment that is not performing its primary function may be subject to administrative neglect and fail to get proper maintenance or required software/hardware upgrades. This is a major risk, defeating the purpose of the whole scheme.

Another advantage of using both routers at the same time is that this approach ensures that the network manager knows when one fails. It is much easier to detect the failure of an active device than of equipment that is in a passive monitoring mode.

2.4 MULTIPLE BACKUPS

The examples we have introduced so far were of the 1-to-1 cardinality; even in the mutual protection case, we had a pair of 1-to-1 configurations. VRRP supports 1-to-N redundancy cardinalities where N can be greater than 1. To illustrate the basic elements of such configurations, let us consider a case in which the R1 router is backed up by two routers, R2 and R3, without any load sharing. Figure 2-6 represents such a configuration.

In this configuration R1 is designated as the default router for all the hosts and the master of the virtual router V1. R2, R3 are pure redundant backups ready to take over the default router role if R1 were to fail. This configuration calls our attention more vividly than the other cases we have studied to the fact that we need a mechanism for deciding which one of the routers, R2 or R3, should become the master if or when R1 fails. Actually, the criterion provided by VRRP for this switchover decision is no different than the one used for the initial decision for mastership. It is based on administratively assigned priorities ranging normally between 1 and 255. The value 0 has a special meaning to indicate that the current master is releasing its mastership responsibility. The higher the number, the higher the priority. At the initiation, the router with the highest priority becomes the master. The same applies to the failover situation in which the protocol elects the router with the highest priority as the master.

The protocol does not, however, prohibit the assignment of the same priority to different routers, and for that reason it specifies another criterion to be used as the tiebreaker in the case of two contending backups with the same priority. This second criterion is based on the specific IP address of a router. Each VRRP router is associated with a set of IP addresses identifying its multiple (real) interfaces. Using an algorithm (the smallest one, the first one in the list, etc.) a router selects one of its IP addresses as its *primary IP address*. VRRP uses this primary IP address as the tiebreaker in deciding between two VRRP routers with the same priority. In such a decision, the router with the greater primary IP address wins.

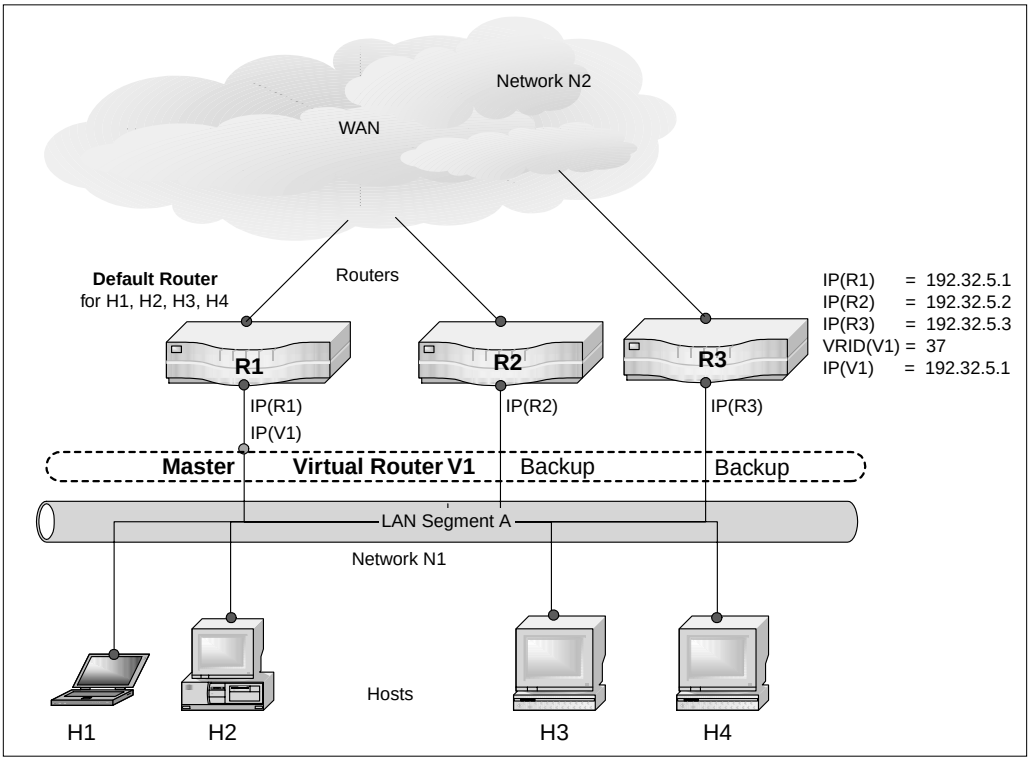

FIGURE 2-6. *A virtual router with multiple backups*

2.5 OWNERSHIP

We have already described different aspects of the mechanism VRRP defines for protecting interfaces, enabling a router to act as a default first hop router. In all the configurations we have discussed so far, the IP address protected by the virtual router was the real IP address of one of the participant VRRP routers. Take the configuration depicted in Figure 2-6. In this configuration the IP address protected by the virtual router V1 is the IP address of the router R1. We refer to such routers, the real IP address(es) of which is associated with the virtual router, as *owner*. In Figure 2-6, by definition, R1 is not only the master but also an owner. For example, in Figure 2-6, the IP address IP(V1) protected by virtual router V1 is also a real IP address IP(R1) of the VRRP router R1. For that reason, R1 qualifies as the owner of the virtual router V1.

Ownership leads to some special characteristics in a VRRP router. VRRP requires the owner to assume the priority 255 and to become the master at the initialization. Another characteristic is related to the situation in which a failed master

becomes operational again. In such cases the owner becomes unconditionally master again. By the same token, as long as it is operational, an owner remains master independent of the status of the other VRRP routers.

The situation is a little bit more complex when neither the current master nor the router that is in the process of becoming operational is an owner. In that case the result depends on the priorities of the routers as well as on the value of a flag called *preemption mode*. The preemption flag regulates whether a new operational nonowner can displace the current master based on its priority.

For the sake of brevity, let us call a master becoming operational after a failure a *new contestant* and use the term *incumbent* for the current master from the perspective of its bid for a new election. As indicated above, since the mastership of an owner is unconditional, a new contestant cannot displace an incumbent owner. By the same token, a new contestant that is also an owner always displaces the incumbent.

But if the incumbent and the new contestant are not owners, in such cases the decision depends on the priorities and the preemption mode. If the priority of the incumbent is higher than that of the new contestant, the incumbent stays in its office. But if the priority of the new contestant is higher than that of the incumbent and if the preemption mode is set to true, the new contestant becomes the master. On the other hand, if the preemption mode is defined as false, in such cases the incumbent keeps its office even if it were to have lower priority. Table 2-2 summarizes the logical structure of this decision flow.

TABLE 2-2. *Impact of Ownership in a New Election*

INCUMBENT	NEW CONTESTANT	PREEMPTION MODE	RESULT
Owner	Not an owner	False	No change
Owner	Not an owner	True	No change
Not an owner	Owner	False	Change
Not an owner	Owner	True	Change
Not an owner	Not an owner	False	No change
Not an owner	Not an owner	True	Requires election

From Table 2-2, you can deduce that there can be at most one master in a given virtual router. For that reason, you do not see any row in which both the incumbent and the new contestant are owners. Note that we did not discuss above the equality cases for priorities. This subject is covered extensively in Chapter 4.

2.6 VIRTUAL ROUTERS WITHOUT OWNER

Now that we have discussed the issues of ownership and nonownership, we may call attention to the assertion that there can be at most one owner in a virtual router.

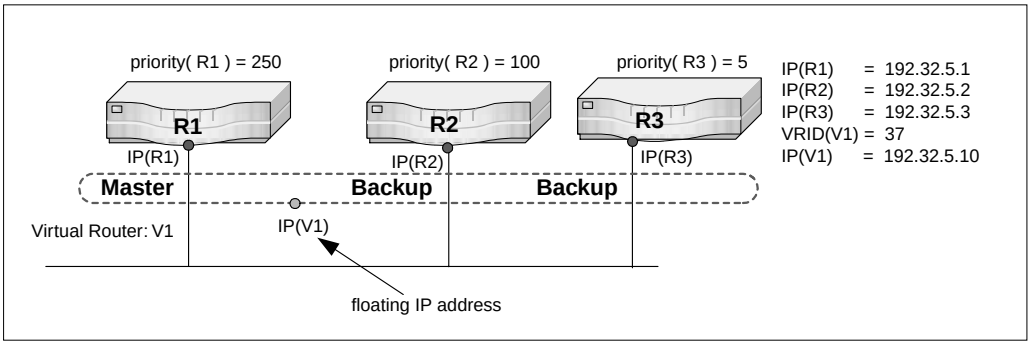

FIGURE 2-7. *Virtual router without an owner*

The term used *at most one* suggests that it is possible to have virtual routers protecting IP addresses that are not owned by any specific router. As a matter of fact, it is not uncommon to have virtual routers in which none of the group members is an owner of the protected address(es). Figure 2-7 depicts such a virtual router.

We have three VRRP routers in Figure 2-7—R1, R2, and R3—all members of the virtual router V1. Since the IP address, IP(V1) protected by the V1 is not the real IP address of any of these routers, we call V1 a virtual router without an owner. We refer to the address(es) such as IP(V1) protected in virtual routers without owners as *floating addresses* or as *pure virtual addresses*.

The main advantage of having a virtual router without an owner is the flexibility it gives to the network administrator. Since the protected IP address is not the real address of any one of the participant routers, the administrator can change these physical routers or their addresses without any need to reconfigure the virtual router itself or the hosts. In Figure 2-7, for example, the network administrator is free to change the IP addresses of R1, R2, or R3 without having any impact on the virtual router V1—that is, of course, as long as these addresses are in the same subnet 192.32.5.0.

2.7 ONE BACKUP PROTECTING TWO MASTERS

Since we have covered 1-to-1 and 1-to-N redundancies, it may be informative to consider how to implement an N-to-1 redundancy using VRRP. To illustrate this case, we install three VRRP routers—R1, R2, and R3—into our local networks consisting of two segments. We would like R2 to back up both R1 and R3. To achieve this, we define two different virtual routers, V1 and V2, and designate R1 as the master of V1 and R3 as the master of V2. The VRRP router R2 assumes the role of the backup both in V1 and V3. Figure 2-8 illustrates this configuration.

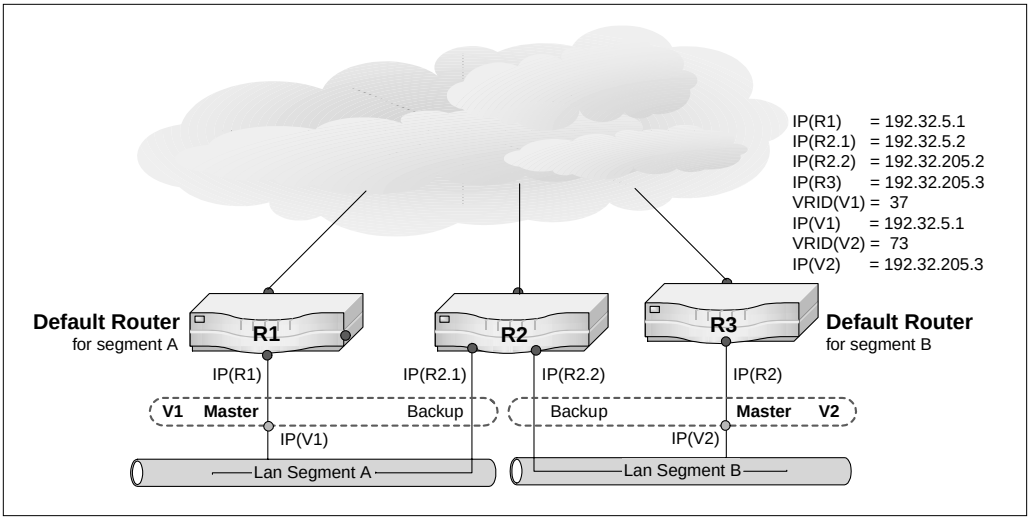

FIGURE 2-8. *One backup for two masters*

Note that in Figure 2-8, at the failure of R1, R2 assumes the responsibility of handling IP(R1), and if during that period R3 also fails, R2 also starts handling packets forwarded toward IP(R2). It goes without saying that this arrangement may lead to unacceptable service degradation unless the network is overengineered and/or R2 is a mighty powerful box.

Also note that the two virtual routers in our illustration are associated with two LAN segments. R1 is on A and R3 is on B, whereas R2 is associated with both through its two interfaces: R2.1 and R2.2.

Another point to consider is that in this configuration the router backing up two masters is not protected. In the event of its failure, both R1 and R2 become unprotected. The very structure of N-to-1 leads to this shortcoming. On the other hand, this configuration is not uncommon and without some merits; given the cost considerations for backup lines, from an economic point of view, N-to-1 arrangements may be quite plausible.

2.8 VIRTUAL ROUTERS WITH MULTIPLE IP ADDRESSES

VRRP Internet Draft hints at the possibility of protecting multiple IP addresses with a single virtual router by using the expression *IP address(es)* without explicitly going into the details of this characteristic. As a matter of fact, a physical interface may have multiple IP addresses that can be on the same (usually rare) or different subnets, the latter being referred to as *multinetting*. Network administrators commonly use

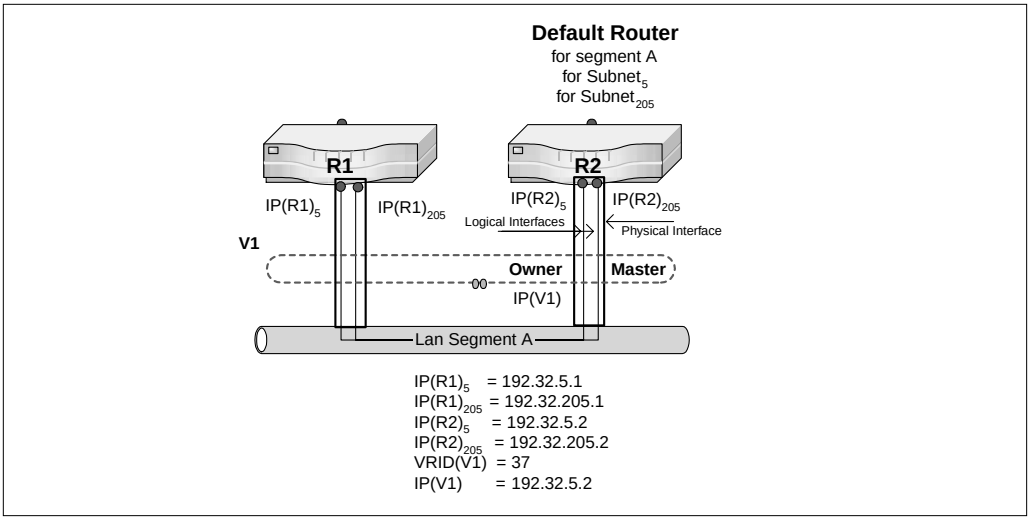

FIGURE 2-9. *A virtual router protecting multiple IP addresses*

multinetting when they need to renumber a network, that is, assign new addresses with a different subnet definition. In such cases, to ease the transition both subnets are maintained for a brief period.

When a VRRP router has more than one address associated with its interface, it can have all its addresses protected by one single virtual router. In such an arrangement, the router in question qualifies as the owner of all addresses, but the router picks one of those addresses as primary to be used for the purposes of VRRP traffic.

These considerations apply independent of whether the protected multiple addresses are on the same subnet or not. The multinetting case is more interesting, though, and demonstrates that VRRP can still be used without significant reconfiguration during the periods of transitions in networks. Figure 2-9 depicts a virtual router protecting a default router configured for multinetting.

Note that in Figure 2-9 we have one LAN segment that is partitioned into two subnets: 192.32.5.0 and 192.32.205.0. Router R2, the owner of both addresses $IP(R2)_5$ and $IP(R2)_{205}$, is the master of the virtual router V1. Router R1 is the backup and is similarly configured, that is, it has been configured for multinetting with an IP address on each subnet like R2. Although the IP addresses are in different subnets, they are associated with the same virtual router V1, and they are both protected by the same virtual router V1. One of the addresses designated via configuration as primary would be picked for V1; in this example, 192.32.5.2 and router R1 would be the backup. The VRRP mechanism and VRRP exchanges will be on this subnet 192.32.5.0, and 192.32.5.2 will be the primary IP address of virtual router V1.

However, the other address, 192.32.205.2, would simply piggyback on the protection offered by VRRP to the primary address because both of them are on the same interface. As long as R2 is operational, R1 will stay in the backup status; but when R2 fails, VRRP will detect the failure and R1 will become the master and route on behalf of 192.32.5.2 as well as 192.32.205.2, since R1 is configured accordingly. In other words, all IP addresses in a multinetted interface will switch over to the backup. We do realize that it is possible to have one virtual router, not multiple ones, to protect more than one IP address as long as all physical routers within a virtual router have a multinetted configuration on the same set of subnets. See Chapter 8 for possible misconfigurations.

2.9 FAILURE CASES

So far we have been discussing different scenarios to illustrate how VRRP's switchover mechanism kicks in under different configurations. It may also be informative to look at different failure points in our context and indicate in which cases VRRP is of help and in which cases tools other than VRRP are required. To list all possible failure points, we start with our R1 router that acts as the master of the V1 virtual router. Figure 2-10 depicts the points of failure on R1.

f0 indicates the failure of the R1 as a whole. Different causes may lead to such a failure: the crash of the operating system, critical hardware failures, somebody accidentally unplugging the device. Whatever the cause, with a failure of f0 category, R1 becomes totally unoperational. f1 corresponds to the failure of the IF(R1.1) interface

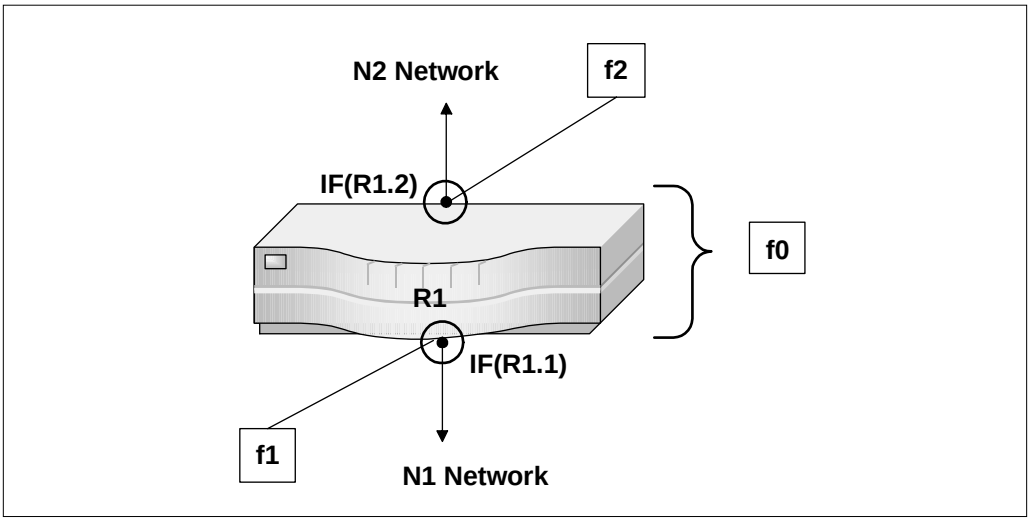

FIGURE 2-10. *Points of failure around a router*

connecting R1 to N1 network. f2 corresponds to the failure of the interface IF(R1.2) connecting R1 to N2 network.

By looking at the larger context of our branch office LAN, we observe three additional failure points. Figure 2-11 establishes such a context in our discussion with VRRP and RIP running in the local network (N1) and BGP4 covering the cloud.

Cloud failure in Figure 2-11 represents a complete failure of the WAN. In such cases, the branch office LAN segment becomes completely isolated. Given the current configuration, neither VRRP nor a dynamic routing protocol would be of help to restore the service. For that reason, we don't include the cloud failure case in our discussion.

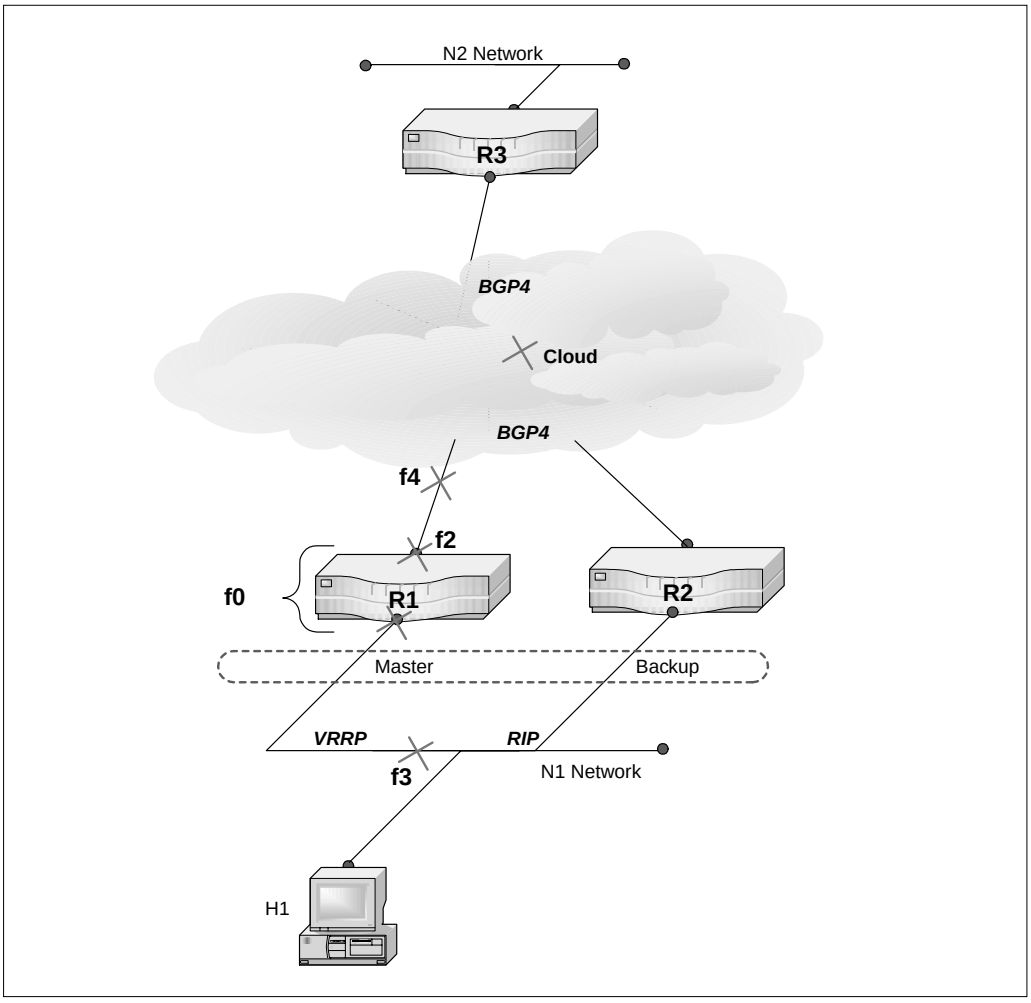

FIGURE 2-11. *Failure cases in context*

f3 represents the failure of the N1 as a whole but with both R1 and R2 staying operational. Finally, f4 represents the case where the IF(R1.2) is operational but a failure occurs in the cloud, making R1 unreachable but keeping R2 connected.

To be more comprehensive, we distinguish the direction of the potential traffic flows. We use the symbol ⇑ to represent the traffic flowing from the H1 host toward the N2 network; we use the symbol ⇓ to represent the traffic in the opposite direction, the traffic flowing from N2 toward H1. In order to contrast the place where VRRP would be of help, and cases where a dynamic routing protocol would be required, we assume that RIP runs on N1 and BGP4 across the cloud between R1, R2, and R3. Note that by using multiple service providers, one creates redundant clouds unless they share resources. This is, of course, an additional (service, organization-oriented) dimension of the availability.

Given five failure points and 2 traffic flow directions, all in all we have ten cases to study. We expect some of them to be equivalent from the perspective of our discussion—relevance of VRRP to the specific case.

The study of Table 2-3 makes clear that VRRP is designed for the protection of a local interface for an outgoing traffic. The switchover kicks in when IF(R1.1) becomes unoperational through f0 or f1. The f4 case is a local catastrophe. Neither VRRP nor dynamic routing protocols are of help in this case. This predicament

TABLE 2-3. *Possible Failure Cases in a Context*

FAILURE	DIRECTION	EFFECTS
f0	⇑	The master is unoperational. R2 becomes the new master, starts forwarding H1 traffic toward N1. **VRRP helps** in this case.
	⇓	The master is unoperational. R2 becomes the new master. VRRP is not of help in this case. R3 detects the failure of R1 through BGP4 and reroutes the traffic to R2.
f2	⇑	The master stays operational, no switch over take place. But detecting the failure of IF(R1.2), R1 through RIP reroutes the traffic to R2.
	⇓	The master stays operational. But R3 detects the failure of R1 through BGP4 and reroutes the traffic to R2. The effect same as f0 .
f1	⇑	The backup in R2 recognizes the unavailabity of the master and performs the switch over. **VRRP helps** in this case.
	⇓	R3 detects N1 as unreachable through BGP4 and reroutes the traffic to R2.
f3	⇑	N1 network is fully isolated. Neither VRRP nor RIP can be of help in this case.
	⇓	N1 network is fully isolated. Neither VRRP nor BGP4 can be of help in this case.
f4	⇑	R1 detects the link failure. RIP reroutes the traffic to R2.
	⇓	R3 detects the failure. and through BGP4 reroutes the traffic to R2.

reminds us that there are other single points of failure in the system that cannot be fixed by just relying on protocols. In cases where IF(R1.1) is not impacted, VRRP is not relevant. To preserve availability in such cases, dynamic routing protocols are handy. Two observations may be appropriate at this juncture: First, dynamic routing protocols are also designed to preserve network availability. This is actually the topic of our discussion in Chapter 1 under the title *Availability at Layer 2 and Layer 3*. Second, VRRP is not a routing protocol. It is a first hop router redundancy, failover protocol, designed for cases in which hosts on a LAN do not run dynamic routing protocols.

This fact reminds us that dynamic routing protocols contribute to the preservation of network availability.

2.10 A FEW WORDS ABOUT ROUTER VIRTUALITY

So far we have been discussing VRRP without talking about the adjective "virtual" in the name of protocol. Yes, we have defined the specific use of the term "virtual router" as it is intended in VRRP terminology. Now that we have an overall understanding of the protocol, we may say a few words about router virtuality in VRRP and other uses.

Encarta World English Dictionary[1] describes the first usage of *virtual* as "being something in effect even if not in reality." The origin of the term goes back to Latin *virtualis* of the 14th century, derived from *virtus* (virtue). From "having the power" comes "so in effect." Encarta also defines *virtual* to mean "generated by computer." We may say that in computer and networking terminology the two usages merge as "creating something so in effect using computing or networking technology." Nowadays we use "virtual" both as a technical term and an ordinary idiom. Examples abound: virtual memory, virtual community, virtual corporation, virtual reality, java virtual machine—just to mention a few.

But specifically the terms virtual router, virtual address, virtual IP address, virtual MAC address are an integral part of our discussion. Here we'll focus on router virtuality.

Assuming reality is physical reality, "virtual" may be expressed as the opposite of physical. But independent of its metaphysical implications, this expression may be misleading, since the term "physical" immediately creates the association with the first layer of the OSI reference model: physical layer and things related to it. The contrast between physical and virtual router may be just fine; contrasting a virtual address with physical address or even a virtual with a physical interface may be very misleading. For that reason, it is safer to use "real" as a contrasting term in our context: real address, real interface, real router.

Actually, the definition "so in effect" works quite well for our purpose: a virtual router is an entity, that is, a router of a certain kind, in effect without being so in reality.

1. Soukhanov, Anne H. *Encarta World English Dictionary* (New York: St. Martin's Press), 1999.

VRRP virtual routers consist of real routers behaving in such a way that in effect they act as if they were one single router, one single default first hop router on a LAN segment. We may also look at a VRRP virtual router as a form of interface virtualization. Figure 2-12 can be used to clarify what we mean by interface virtualization.

In Figure 2-12, we have two real routers, R1 and R2. R1 has three interfaces, RR1.1, RR1.2, and RR1.3, whereas R2 has two interfaces, RR2.1 and R2.2. Here we use the notation RR.X.Y to articulate that the interface we are referring to is a real interface of a real router. R1 and R2 are the members of a VRRP virtual router to protect the interface R1.1. In this sense, we can say that R1.1 is the interface of the virtual router, VR1. The IP address(es) of the R1.1 is/are also the IP address(es) of the VR1. For that reason, we use the term *virtual IP address* to refer to IP(VR1) or IP(RR1.1), since it is mapped to the interface of a virtual router. In our configuration note that R2 participates in VR1 only on its interface R2.1, not on the other interface R2.2, and the same exclusion applies to R1.2 and R1.3. With all these interface virtualizations or mappings, VRRP multiple routers impersonate one single router.

There is another sense of the term "virtual" in "virtual router," the discussion of which can help us to relate better to the concept of VRRP virtual router. The term we have in mind is "multiple virtual routers" or MVRs. MVR refers to the multiple instances of the routing function within a single device. If the operating system is a

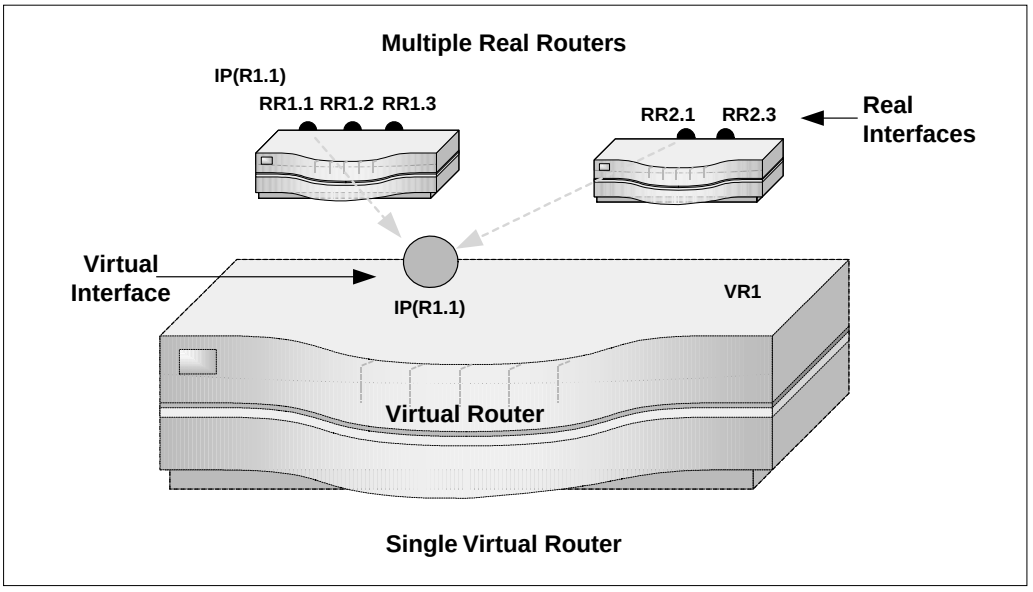

FIGURE 2-12. *Single VRRP virtual router consisting of two real routers*

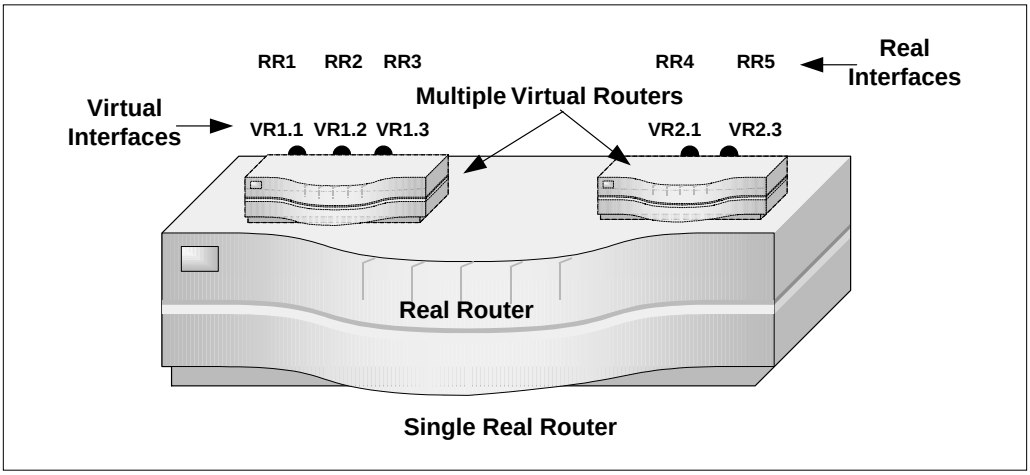

FIGURE 2-13. *Multiple virtual routers within one single real router*

multitasking OS, different instances may be implemented—for example, as different tasks. Figure 2-13 represents a "virtual router" in this second sense.

Figure 2-13 shows one real single router hosting two virtual routers. In this example, we see a mapping of the real interfaces to the virtual interfaces. We allocate real interfaces—RR1, RR2, and RR3—to the virtual router on the left. We assign RR4 and RR5 to the virtual router on the right. These virtual routers operate as if they were distinct routers. They forward packets among their interfaces and collaborate with each other as if they were different devices. It's likely that they maintain different route tables. They may even be running different routing protocols. Another variation on this term is related to routing domains. In some devices handling different routing domains, the term "virtual router" is used as if a distinct router were dedicated to each routing domain.

Since routing touches both forwarding and control aspects of networking, we call attention to the fact that "virtual routing" in the "multiple virtual router" sense may also be touching both aspects. It is conceivable that MVRs constitute different control planes but share a common, albeit distributed, forwarding plane. It is not out of the question for MVRs also to have their dedicated forwarding engines.

To summarize: MVRs are different instances of the routing function within one single network element. Their role is to impersonate the control and in some cases the forwarding aspects of a distinct router. In the case of VRRP virtual router, we deal with multiple real routers collectively impersonating one single router, which they do in terms of acting as the default first hop router in a LAN segment.

2.11 CASE STUDIES

So far we defined and clarified VRRP concepts using simple configurations to keep the focus on the understanding of the protocol. Now that we have covered enough ground as an overview, we consider some typical VRRP deployment cases and call attention to some economic, risk-management aspect of the specific configuration. To do that we'll use different portions of the enterprise network we introduced at the beginning of this chapter in Figure 2-1. For the sake of facilitating the discussions, we use some made-up product names (MRM 100, MRM 5000, etc.) to create a context in which we can cover the trade-off considerations within a product line or a product category. With the same motivation, we also attribute some features to the products under discussion (support for dial-on-demand, filtering capabilities, etc.) so that the rationales for the described configurations become clearer.

2.11.1 Dial Backup and No Load Sharing

The first standard configuration in the branch offices of the enterprise network under study consists of three routers: one high-end model and two low-end models of the same product line. The high-end model is connected to the cloud through a Data Service Unit/Customer Service Unit (DSU/CSU) using leased lines: T1 or T3. The *router* MRM 5000 shown in Figure 2-14 represents such a configuration.

MRM5000 is the gateway of the branch office to the cloud and through the cloud to the Corporate Center. If it were not backed up, the failure of MRM 5000 would cause the branch office to lose all its network connectivity to the external world and thus to become almost unoperational, since all parts numbers and sales prices are stored in the Corporate Data Center. The economics of the company do not permit the purchase of another high-end router in addition to the lease of another T1 line for typical branch offices. For that reason, two MRM 100s, the lower-end models of the same product line with built-in dial backup modems, are used as backup equipment. Given that MRM 100s are lower-end models, their reliability and internal redundancy features may not be as strong as the higher-end models. For that reason, using not one but two backups with a 1-to-2 cardinality appears to be a sensible risk-management strategy. Of course, this configuration is most effective only in conjunction with an availability software such as VRRP. With some commercial products, VRRP may come as a feature of the standard software. In some others, it may require a special purchase or be part of a special software bundle. In the configuration discussed, we define MRM 5000 as the default router for all the workstations, PCs and laptops in the office, that is, all traffic that cannot be routed directly gets directed to MRM 5000 as long as it is operational. We configure the MRM 100s as two backups of MRM 5000, one with priority 5, the other with priority 250, to avoid any race condition that may occur because of closeness of priorities. MRM

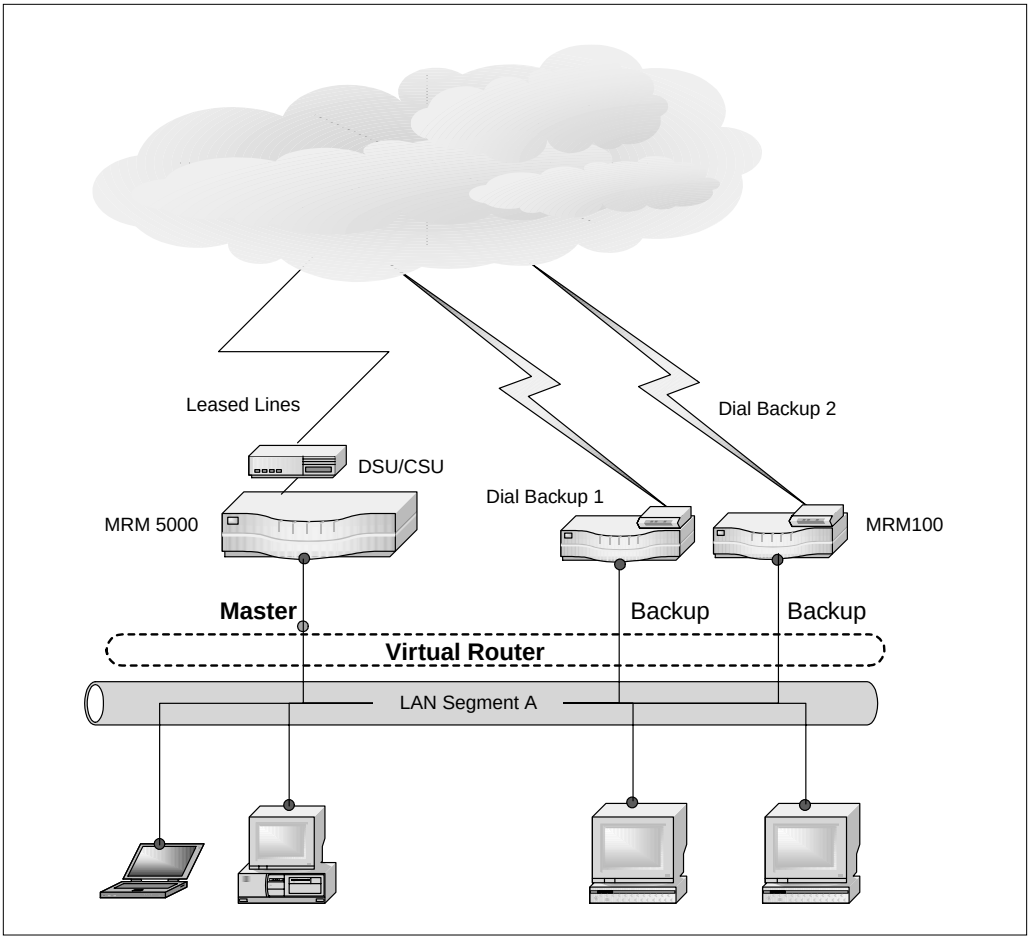

FIGURE 2-14. *Dial backup without load sharing*

100's dial backup features are also configured for dial-on-demand, so that when the router becomes the master, a trigger event activates the modem to dial into the modem bank of the Corporate Office. Although MRM 100 supports a simple agent that can detect the state change in VRRP and activate the modem when the router becomes the master, we prefer the use of the dial-on-demand feature of the MRM 100 to handle also cases that are not related to VRRP failover situations. With this feature enabled, the presence of the packets in the router destined toward the cloud activates the modem.

MRM 100s have built-in policy agents that can be configured to suppress a specified type of traffic. This is particularly useful to suppress nonessential traffic

while the MRM 100s are acting as masters since they are connected over a low-bandwidth dial-up link. For example, when the MRM 5000 fails and MRM 100 takes over as the master, the policy filters could be set up only to pass through essential traffic between corporate network and the home network such as intranet network, database, and firewall server connectivity. It is good practice to program the filters to suppress email and Internet/Web traffic to alleviate the degradation of service after the VRRP switchover.

2.11.2 Load Sharing in the Branch

Some branches of the enterprise network have different networking and availability requirements. Because of their special criticality to the business and their special impact on the bottom line, different budgets are allocated to such organizations. If we were to call the configuration we discuss in Figure 2-14 *small branch network, big branch network* would be a proper label for the solution depicted in Figure 2-15. This configuration consists of two MRM5005s connected to the cloud through leased lines and mutually backing each other up by establishing two VRRP virtual routers. Since MRM 5005s have built-in DSU/CSU capability, we don't represent this equipment explicitly in Figure 2-15.

In the big branch configuration, we define two virtual routers and assign one MRM 5005s as master to a virtual router respectively. Keeping in mind the large amount of computer equipment in the office, we place intelligent hubs BSM 10 on the LAN. We group the default router assignments accordingly. We assign MRM 5005 on the left as the default for the hosts attached to the BSM 10 on the left, and we designate MRM 5005 on the right as the default router for the hosts attached to the BSM 100 on the right.

When both MRM 5005s are operational, they act as the master of their virtual routers and share the burden of handling the traffic originating from the office network. When one of the MRM 5005s fails, the other becomes the default router for all the hosts in the network. Given the hardware-based forwarding engines of the routers and the relatively higher capacity of the leased lines (T3), we assume that the service degradation should not be noticeable. Given this assumption, we enable the policy agents conditionally to suppress nonessential messages only when the traffic demands a certain level of bandwidth.

Although there might be some concerns about the impact of the BSM 100 devices on availability, given the simplicity of these devices, they have lesser chances of being additional points of failure in the network.

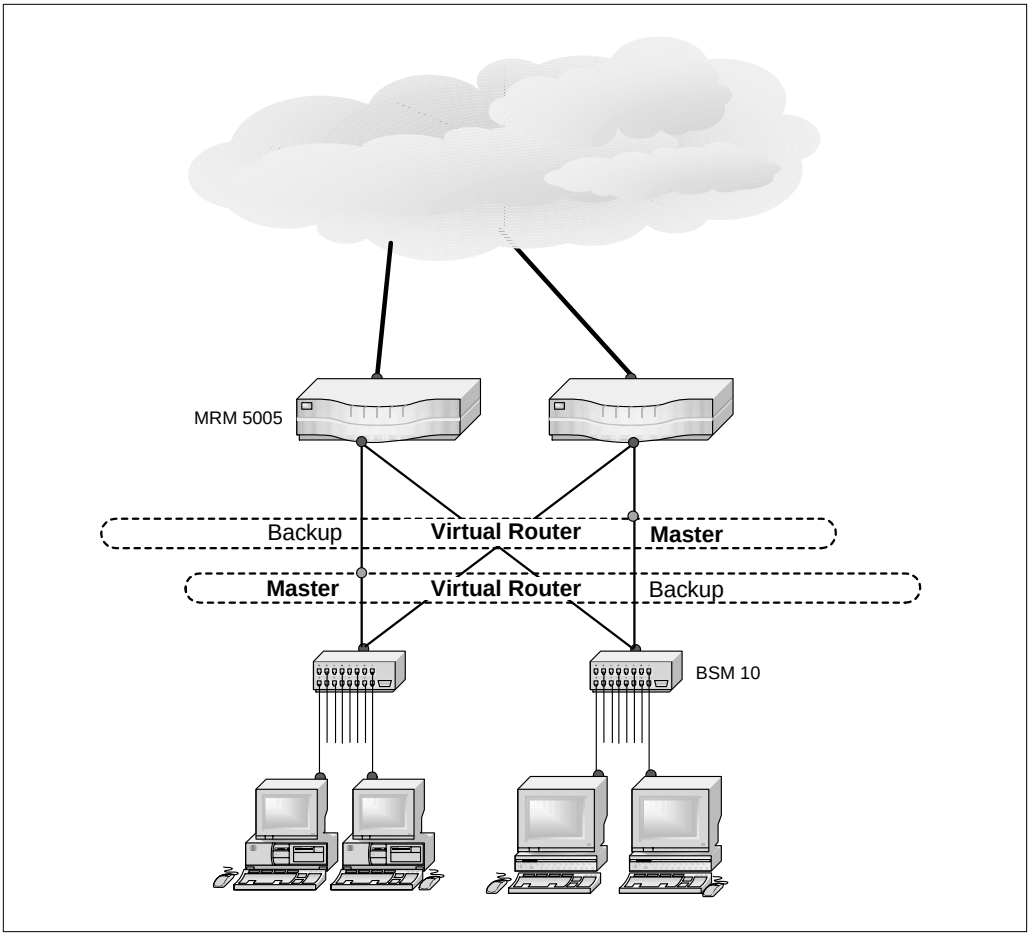

FIGURE 2-15. *Load sharing in a branch*

Summary

This chapter provided an overview of the VRRP and explained its basic concepts by discussing some selected configurations. We learned how to achieve load sharing or load balancing using VRRP; we also learned the technique for not compromising service quality in case of VRRP switchovers. The multiple backups gave us an opportunity to understand better the mechanism of VRRP. Our discussion also touched on different redundancy setups with different cardinalities. We covered how to configure VRRP routers for creating 1-to-1, 1-to-many, and many-to-1 redundancies.

We used stylized configurations to introduce the basic VRRP notions. Through the discussions of these configurations, we have clarified the meaning of terms such as VRRP router, virtual router, master, backup, associated IP addresses, primary IP address, and owner. They are presented here again in summary:

Default gateway or *default router:* A router, the IP address of which (one of its interfaces) is entered in the routing table of the selected hosts on a LAN segment. Thus, these hosts forward all packets addressed to destinations outside the local network to this router. We also use the term *default first hop* in this sense.

VRRP router: a router running VRRP

A *virtual router:* a group of VRRP routers collaborating with each other to establish a failover mechanism to protect the default gateway function. A set of IP addresses is associated with a virtual router, representing the default router for some hosts on a LAN segment.

Virtual Router ID (VRID)*:* a label (an integer) used to name a virtual router.

Master: a VRRP router in a given virtual router that acts as the default router for the hosts on an associated LAN. Master router has the responsibility of responding to the ARP requests related to the IP addresses of the virtual router. This characteristic also entails the responsibility to forward the packets forwarded to the virtual MAC address of the virtual router.

Backup: VRRP routers in a given virtual router that are configured to take over the master role in case of the failure of the current master. This takeover is regulated by the election mechanism specified by the protocol, VRRP.

Load sharing or *load balancing:* Technique of distributing the traffic to different network devices so that we don't overburden one or a few of them, so that we take advantage of all resources.

Priority: An integer representing the order in which backup routers may assume mastership responsibility. The value zero indicates that the current master ceased to participitate in VRRP. The values between 1 and 255 are used in ranking the backups, 100 being the default. The value 255 indicates the ownership. See *owner*.

Primary IP address: An address selected from the IP addresses of a VRRP router. The comparison of IP addresses breaks the tie between backups with the same priority in the election of the master.

Owner: The VRRP router that has the virtual router's IP address(es) as real interfaces address(es). The owner, when functional, responds to the packets addressed to these IP addresses through ICMP pings, TCP connections, and so on.

Preemption mode: A flag regulating whether a new operational, nonowner can displace the current master based on its priority. When the flag is set to false, a nonowner router that becomes master cannot replace the current master even if its priority were to be higher. This replacement is permitted when this flag is set to true.

After considering different failure points in the context of our discussion, we differentiated between the ones covered by VRRP and the ones for which VRRP was of no help. A short excursion on the semantics of "virtuality" to us in the realm of multiple virtual routers gives us another occasion to relate better to the protocol under study.

Finally, equipped with basic concepts and enough familiarity with VRRP's behavior under different configurations, we reviewed some deployment use cases to get closer to the real-world use of VRRP.