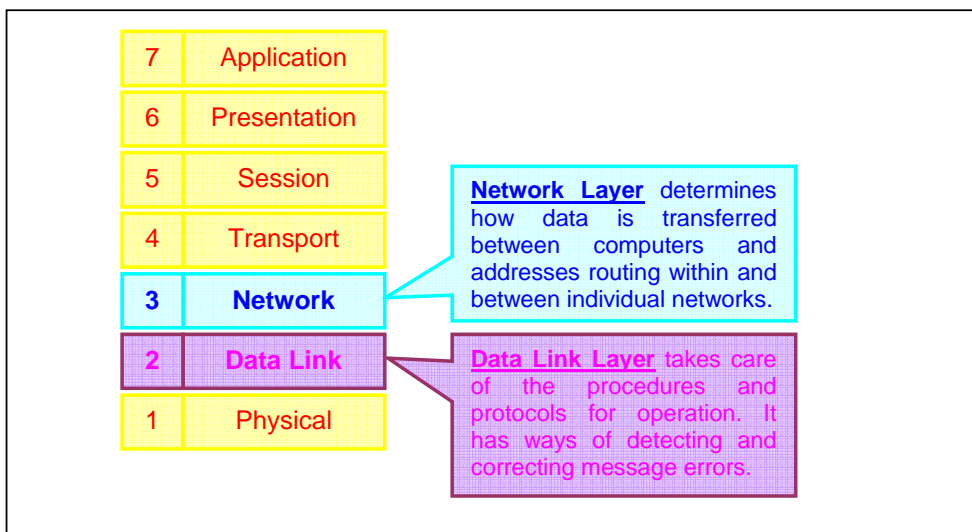


# Layer 3 Switching

## Introduction to Layer 3 Switching

Layer 3 is the Network Layer of the Open Systems Interconnection (OSI) seven-layer model of networking. The Network Layer controls the routing of messages across networks, network flow, and traffic management. It is concerned with knowing network addresses of neighboring nodes, selecting best routing paths, identifying and forwarding of messages. Layer 3 switches combine the functionality associated with traditional routers and switches into the mechanism of a switch.

Traditional switches operate at the Data Link Layer (Level 2) of the OSI model, which controls the flow of data between nodes. Traditional routers are software-based and operate at the Network Layer (Level 3). They used to be core components of enterprise networks, but are relatively slower in performance and have difficulties migrating to newer generation networks. Layer 3 switches offer full switching capabilities and fast routing near switching speeds using silicon-based solutions rather than software routing. In addition, Layer 3 switches also provide secure and more reliable routing.



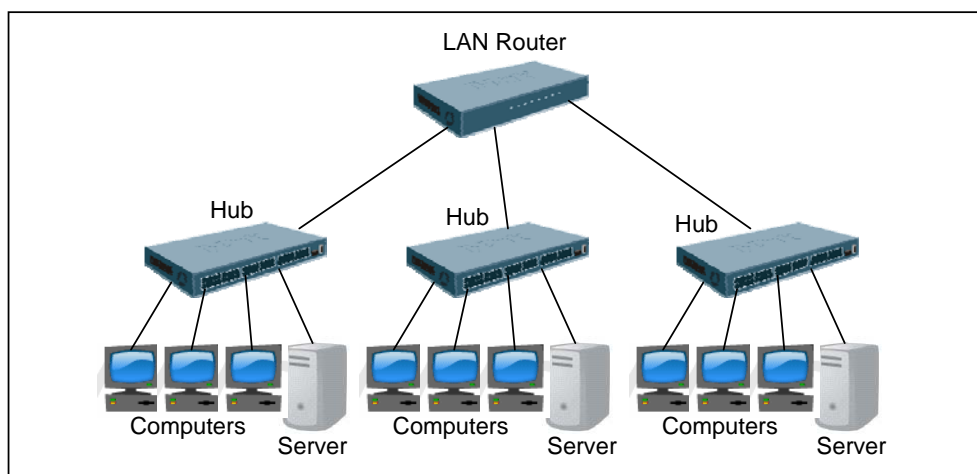
**Figure 1.** Layer 2 and 3 Switches in OSI Seven-Layer Model

A Layer 3 switch is also called IP switch because it is primarily used for routing Internet Protocols (IP). Like traditional router, it can perform packet-by-packet and

flow-based routing. It is able to determine routing paths based on Layer 3 information. Layer 3 switches are ideal for use in network backbone and in a segment providing connectivity of other segments.

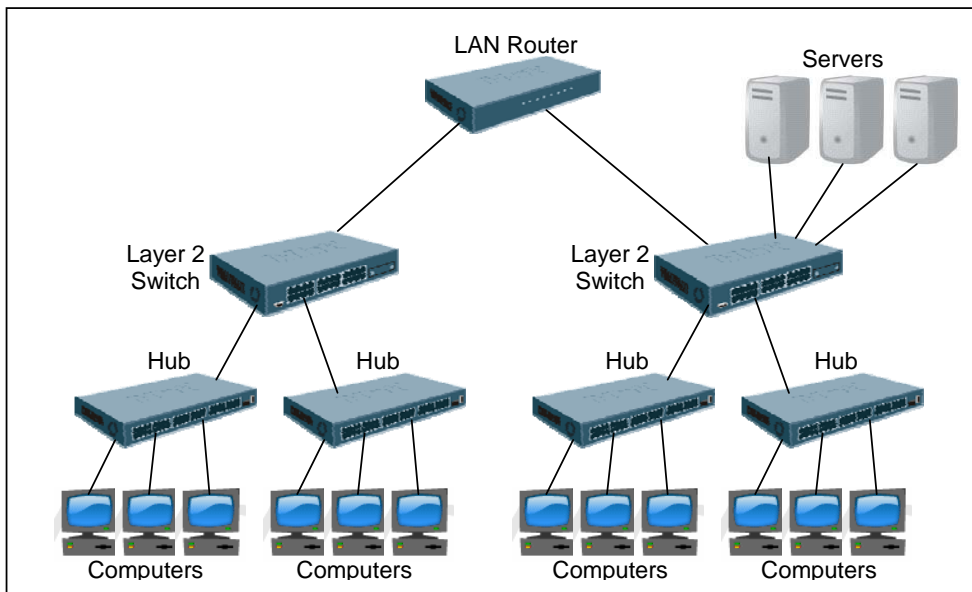
### Need for Layer 3 Switching

When networking was first introduced, it used to be small and flat with peer-to-peer connections on a cable. When bridges were later introduced to expand and connect small networks into larger networks, networking becomes busier and more complex. As the size of networks continues to grow, routers were introduced for use as interconnection devices. Routers are able to provide segmentation and logical structure to networks.



**Figure 2.** Using Traditional LAN Router

Routers, however, are slow and expensive. They are also complex in configuration and management. As demands for performance increase, switches are introduced. Layer 3 switches allow more routing layers within a network and provide faster, simpler, and more cost-effective solutions.



**Figure 3.** Using Layer 2 Switches

Today, with the increase in Internet and Intranet communications, non-local traffic demands on networks have grown beyond the capability of the LAN routers to handle. To work around router constraints, network managers and administrators are forced to segment existing networks into more switched segments. This approach does not resolve the issues in the long run as loading on backbone routers continues to increase.

A better solution to LAN router bottleneck is to use Layer 3 switches, which incorporate both routing and switching technologies into a single device. Similar to the router, Layer 3 switches provide logical segmentation to the network and route data between them. In addition, the Layer 3 switches also provide full Layer 2 switching capabilities. With the combined technologies built into Layer 3 switches, companies can benefit from greater network performance overall.

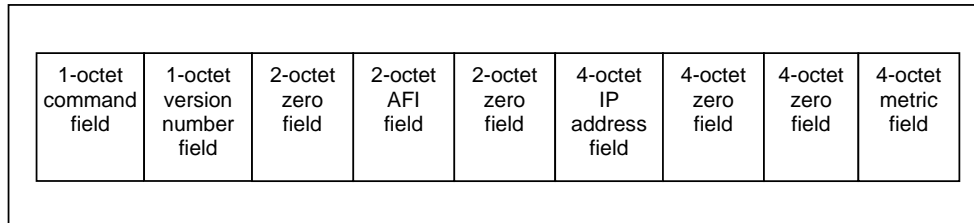
### **Routing Information Protocol**

Routing Information Protocol (RIP) is a widely used routing protocol. Routers or switches with RIP enabled can construct a routing table that lists all the other hosts it knows. This is done by using RIP to send the entire routing table to the neighbor host, which in turn passes the information on to its next neighbor and so on until all hosts within the network have the same knowledge of routing paths. The routing switch or

router will compare the routes to the destination address, identifies the lowest-cost path, and updates the routing table to reflect the new route.

### RIP Packet Format

An IP RIP packet consists of nine fields. The following is a summarized description of the IP RIP packet format fields:



**Figure 4.** IP RIP Packet Consists of Nine Fields

- **Command**

This field indicates whether the packet is a request or a response. The request asks a router to send all or part of its routing table. The response can be an unsolicited regular routing update or a reply to a request.

- **Version Number**

This specifies the RIP version used and can signal different incompatible versions.

- **Zero**

This is added to provide backward compatibility with pre-standard varieties of RIP. The default value is zero.

- **Address-Family Identifier (AFI)**

This specifies the address family used. RIP is designed to carry routing information for several different protocols. Each entry has an address-family identifier to indicate the type of address being specified. The AFI for IP is 2.

- **Address**

This specifies the IP address for the entry.

- **Metric**

This indicates the number of internetwork hops traversed in the trip to destination. The value in this field is between 1 and 15 for a valid route, or 16 for an unreachable route.

### RIP Routing Table

The RIP routing table has five fields.

- **Destination IP Network Address**

This is the destination of the data packet that the router receives. The router will look up the destination network address in the routing table to determine where to send the packet.

- **Metric**

This is the value of the route from the source to the destination. It contains the sum of the costs associated with the end-to-end path across the network from the router. Each RIP link cost is 1. The maximum cost, or metric, is 16.

- **Next Hop IP address**

This is the IP address of the next router in the path to the end destination. This field will exist if the destination IP address is not connected directly to the router.

- **Route Change**

This field is not always implemented by router manufacturers. It is used for identifying changes in routes to specific destination entries.

- **Route Timers**

There are three timers that are associated with each route:

- Update Timer

This timer initiates routing updates.

Deleted: that

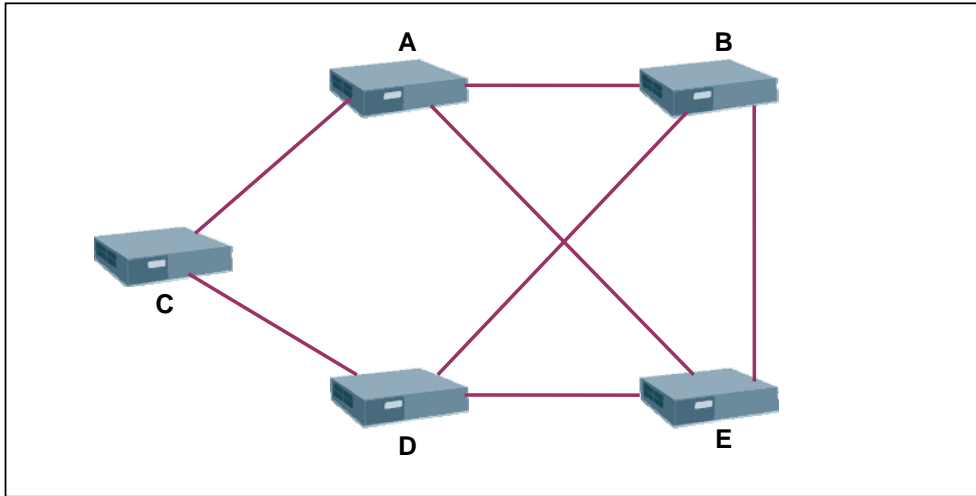
- Route Timeout

This is the amount of time a route may be without update before the route is marked as invalid.

- Route-Flush

This is the amount of time after the route timeout timer has expired before the route is flushed from the routing table.

The network topology shown below is an example illustrating the contents of a RIP routing table. It is focused on the hop count.



**Figure 5.** Example to illustrate RIP

Router	Destination Hop	Next Hop	No. of Hops	Network Path	
<b>A</b>	B Network	B	1	Directly Connected	
	C Network	C	1	Directly Connected	
	D Network	B	2	B-D	
		C	2	C-D	
		E	2	E-D	
		B	3	B-E-D	
		E	3	E-B-D	
	E Network	E	1	Directly Connected	
	<b>B</b>	A Network	A	1	Directly Connected
		C Network	A	2	A-C
		D	2	D-C	
		E	3	E-A-C	
		E	3	E-D-C	
D Network		D	1	Directly Connected	
E Network		E	1	Directly Connected	
<b>C</b>		A Network	A	1	Directly Connected
	B Network	A	2	A-B	
		D	2	D-B	
		A	3	A-E-B	
		D	3	D-E-B	
	D Network	D	1	Directly Connected	
	E Network	A	2	A-E	
		D	2	D-E	
		A	3	A-B-E	
		D	3	D-B-E	
<b>D</b>	A Network	B	2	B-A	
		C	2	C-A	
		E	2	E-A	
		B	3	B-E-A	
		E	3	E-B-A	
	B Network	B	1	Directly Connected	
	C Network	C	1	Directly Connected	
	E Network	E	1	Directly Connected	
	<b>E</b>	A Network	A	1	Directly Connected
		B-Network	B	1	Directly Connected
C-Network		A	2	A-C	
		D	2	D-C	
		B	3	B-A-C	
		B	3	C-D-B	
D-Network		D	1	Directly Connected	

Reference Only

Formatted: Font: Italic

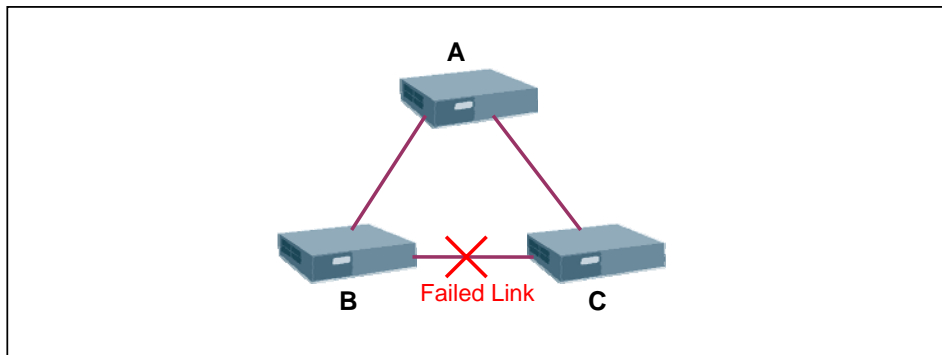
## RIP Routing Convergence

The routers must be able to converge when the network topology changes. Convergence is the mechanism that all routers agree with each other on what the new network topology looks like. When there is a link failure, each interconnected router will converge and update its routing tables. The same process will occur when a router fails.

The following are several mechanisms that will affect the convergence of the RIP routers:

### - Count to infinity

The network shown below is an example of a failed link between Router B and Router C.



**Figure 6.** RIP Network with a Failed Link

Based on this example, the RIP Routing Table will show the following before the network failure:

Router	Destination Hop	Next Hop	No. of Hops	Network Path
<b>A</b>	B Network	B	1	Directly Connected
		C	2	C-B
	C Network	C	1	Directly Connected
		B	2	B-C
<b>B</b>	A Network	A	1	Directly Connected
		C	2	C-A
	C Network	C	1	Directly Connected
		A	2	A-C
<b>C</b>	A Network	A	1	Directly Connected
		B	2	B-A
	B Network	B	1	Directly Connected
		A	2	A-B

When the connection between Router B and Router C fails, a few events will take place. Router B will try to connect to Router C, but it does not have a direct connection. Router B learns that Router A has a connection to Router C, but Router A announces that Router C can also be reached directly via Router B. Router B in turn announces that it can reach Router C through Router A. This will go back and forth between the two routes, and add 1 to the hop count until the next routing update takes place. The routing update default is 180 seconds or infinity.

The following table shows the routing tables after all the routers have counted to infinity to determine that each node is reachable after the link failed:

Router	Destination Hop	Next Hop	No. of Hops	Network Path
<b>A</b>	B Network	B	1	Directly Connected
		C	16	Unreachable
	C Network	C	1	Directly Connected
		B	16	Unreachable
<b>B</b>	A Network	A	1	Directly Connected
		C	2	C-A
	C Network	C	1	Directly Connected
		A	2	A-C
<b>C</b>	A Network	A	1	Directly Connected
		B	<del>16</del>	<del>Unreachable</del>
	B Network	B	16	Unreachable
		A	2	A-B

Deleted: 2  
Deleted: B-A

This above scenario describes how a link failure table handling can pose an issue pertaining to the amount of time for the 16-hop count to be achieved. The datagram traffic circles around the two nodes and will never be able to reach its ultimate destination until the network converges upon the next routing update.

Two methods that may be used to avoid the count to infinity problem are split horizon and triggered updates.

**- Split Horizon**

With split horizon, the router does not announce a route to the same interface from which it was learned. It is designed to prevent routing loops in a network. However, the router must wait for the destination to be marked unreachable after a route has timed out and flushed from the table. This process takes six update messages of total 3 minutes before each routing table is updated with the inactive link. During this time, five update intervals can pass and each router can misinform another as to whether certain destinations are reachable. Split horizon with poison reverse can address and solve this problem.

**- Split Horizon with Poison Reverse**

Split horizon with poison reverse is more effective than split horizon because the six update cycles do not have to pass, to stop a routing loop. It takes a more proactive stance in managing and updating the routing tables. When an inactive

link is detected, RIP with poison reverse sets the metric for that destination to infinity for the next routing update.

Although split horizon with poison reverse is preferred over the traditional split horizon, large networks with multiple paths are still concern because RIP is still subject to the counting to infinity problem of routing updates. Triggered updates were introduced to solve routing loops problem that is caused by the “counting to infinity” operations.

#### - **Triggered Updates**

Triggered updates are used to speed up convergence of a RIP-routed network. Triggered updates are rules in the routing protocol that require routers to immediately broadcast an update message whenever the route metric changes, without waiting for the next 30-second regular update interval to pass.

Although triggered updates have significant improvement over split horizon and poison reverse, the question of whether each router in the network can receive and update its tables in reasonable time is still an issue. Hold-down timers solve this problem by working with triggered updates.

#### - **Hold-Down Timers**

Hold-down timers are used for starting a clock count down when a triggered update is sent. When this hold-down timer hits zero, the router will not accept neighbor updates for the route in question.

The use of a hold-down timer prevents a RIP router from accepting and converging on updates for a route that has been invalidated over a period of time. Hold-down timers prevent a router from believing that another router has a path to an invalid destination.

==End of Reference==

Formatted: Font: Not Bold, Italic

### **RIP Concerns and Limitations**

Although RIP has been around for a long time, it still has limitations and concerns that need to be addressed. They are:

- **RIP Hop Count Limitation**

RIP is designed to support networks that are relatively small. When the data packets are forwarded across a router, the hop counters are incremented by the cost of the link over which they traverse. The default is one. When the hop counter hits 15 and the packets are not at its intended destination, it will be dropped, and the destination is considered to be unreachable. This is because the next hop is 16.

- **RIP Fixed Metrics**

RIP is unable to update its cost metrics to adapt to changes in the network topology in real-time. RIP metrics can only be changed manually and are static for the duration. These fixed and static metrics cause RIP to be unsuitable in supporting real-time applications.

- **RIP Network Bandwidth Consumption for Routing Table Updates**

Instead of sending only the updates of the affected route, RIP routers broadcast their entire routing table out to every RIP-enabled interface every 30 seconds, with the exception of split-horizon routes. This can consume large amount of network bandwidth which may be used for carrying data traffic for large networks.

- **RIP Slow Convergence**

RIP routing updates are sent every 30 seconds, and in networking, many things can happen during that timeframe. Furthermore, each RIP router takes up to 180 seconds to converge and invalidate a route. As the network topology grows, this convergence time can increase to an unmanageable state.

- **RIP is Lack of Dynamic Load Balancing**

RIP is not capable of dynamically load-balancing across two or more links. If RIP initially learns a slower path to a destination and then a faster path to the same destination in a later update, RIP will continue to use the slower path because it takes what has first been learned. It will only use the faster path when the slower one is unreachable.

**Summary: Routing Internet Protocol**

RIP is designed for small networks with static configurations and stable links. It does not support dynamic load balancing and it is slow in converging. It consumes bandwidth while routing updates. Other issues include fixed cost metrics and limited hop count. RIP is not suitable for large networks and real-time network applications.

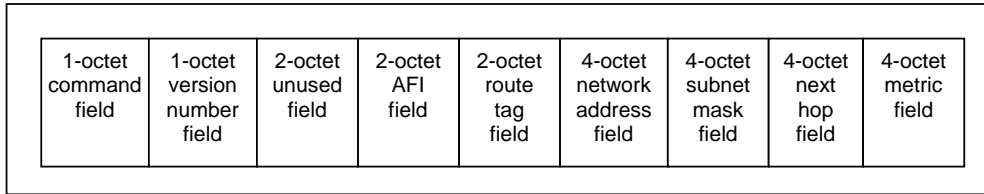
**Routing Internet Protocol Version 2**

Routing Internet Protocol Version 2 (RIPv2) is backward compatible with RIPv1. While RIPv2 shares the same basic algorithms as RIPv1, it supports several new features such as external route tags, subnet masks, next hop addresses, and authentication.

**RIPv2 Packet Format**

RIPv2 allows more information to be included in RIP packets and provides a simple authentication mechanism that is not available in RIPv1.

The following is a description of the IP RIPv2 packet format fields:



**Figure 7.** IP RIPv2 Packet

- **Command**  
This field is the same as RIPv1 which indicates whether the packet is a request or a response.
- **Version Number**  
This specifies the RIP version used. This value is set to 2 for any RIPv2 fields or when authentication is used. Like RIPv1, it can signal the different incompatible versions.

- **Unused**

This has a value set to zero. Its purpose is for RIP backward compatibility.

- **Address-Family Identifier (AFI)**

This specifies the address family used. RIP is designed to carry routing information for several different protocols. It is identical to the RIP's AFI field, with one exception. If the AFI for the first entry in the message is 0xFFFF, the remainder of the entry contains authentication information. Currently, the only authentication type is simple password.

- **Route Tag**

This provides a method for distinguishing between internal routes that are learned by RIP and external routes that are learned from other protocols.

- **IP Address**

This specifies the IP address for the entry.

- **Subnet mask**

This contains the subnet mask for the entry. If it is zero, no subnet mask is specified for the entry.

- **Next Hop**

This indicates the IP address of the next hop to which packets for the entry should be forwarded.

- **Metric**

This indicates the number of internetwork hops that is traversed in the trip to the destination. The value is between 1 and 15 for a valid route, or 16 for an unreachable route.

Additional features of RIPv2 include:

- **RIPv2 Authentication**

RIPv2 authentication is used to authenticate routing response messages that are propagated throughout the network. Authentication of these routing response messages prevents the routing tables from being corrupted by routes from fraudulent sources.

Authentication is achieved using a 16-octet maximum password with no encryption. RIPv2 authentication messages are susceptible to attack by anyone with direct access to the network.

- **RIPv2 Subnet Mask Support (CIDR)**

The subnet mask of the destination IP address is added behind the packet's IP address. It enables RIPv2 to route to a specific subnet.

- **RIPv2 Next Hop IP Addresses Identification**

This next hop identification provides RIPv2 greater efficiency than RIPv1. It prevents unnecessary hops between endpoints in the network. In a pure RIPv2 only network, this next hop identification does not add any significant value, however, if RIPv2 is implemented with other routing protocols, next hop identification is important because without it some routes may not be discovered.

- **RIPv2 Message Multicasting**

RIPv2 multicasting enables the simultaneous delivery of routing table updates to multiple neighbors rather than repeat unicast routing update to each neighbor on an individual basis. RIPv2 multicasting can implement filters to prevent RIPv2 routing updates from being received by RIPv1 routers.

**Differences between RIPv2 and RIPv1**

RIPv2 has several enhancements over RIPv1. These include:

- **Subnet Mask**

RIPv2 carries a subnet mask, while RIPv1 does not. This means RIPv1 will require all router interfaces to be configured with IP addresses in the same network class with the same subnet mask. Subnet mask information can therefore only be derived from the router's configuration.

- **Aggregated Links**

RIPv2 does not require that routes be aggregated on the network boundary, while RIPv1 requires it when routes are sent out over an interface of a different class.

- **Password Authentication**

RIPv2 also supports password authentication security, which RIPv1 does not support

- **Hop Address**

RIPv2 can specify the next hop address, which RIPv1 cannot.

### **RIPv2 Concerns and Limitations**

Although RIPv2 has made several improvements over RIPv1, some concerns and limitations regarding operation remain:

- **RIPv2 Hop Count Limitation**

Like RIPv1, RIPv2 inherited the same maximum hop count limit of 15. This is to maintain backward compatibility with RIP.

- **RIPv2 Count to Infinity**

Like RIPv1, RIPv2 relies on the count to infinity mechanism to resolve certain network error conditions. Routing loops are allowed a relatively long period of time before the “loops” are detected by the hop count, and the route is marked unreachable.

- **RIPv2 Static Metrics for Route Calculation**

Like RIPv1, RIPv2 selects routes based on a static cost metric or hop count. The metric will remain static until it is manually changed.

- **RIPv2 Lack of Alternative Routing Support**

Like RIPv1, RIPv2 maintains only a single route to a specific destination in its routing tables, providing no support for dynamic load balancing. If the “known” route fails, RIPv2 must wait for another routing update to determine the next optimal path to a destination.

### **Summary: Routing Internet Protocol Version 2**

RIPv2 is better than RIPv1 because new features like simple authentication, route tag, and subnet mask are available. However, many limitations from RIPv1 are still inherent in RIPv2. Similar to the earlier RIP, RIPv2 maintains only a single route to a specific destination in its routing tables, providing no support for dynamic load balancing. If the “known” route fails, RIPv2 must wait for another routing update to determine the next optimal path to a destination.

## **Open Shortest Path First**

Open Shortest Path First (OSPF) is a routing protocol developed for Internet Protocol (IP) networks. It is a routing protocol designed for larger and more complex networks that RIP cannot support.

### **Characteristics of OSPF**

OSPF has two primary characteristics. The first is the protocol is open, hence, its specification is for public domain. The second is OSPF is based on the Shortest Path First (SPF) algorithm, also known as Dijkstra algorithm, named after its creator.

OSPF was introduced to overcome some of the following limitations in RIP:

- RIPv1 and RIPv2 have a limit of 15 hops. Any RIP network that spans more than 15 hops is considered unreachable.
- RIPv1 and RIPv2 broadcast full routing table periodically. This consumes large amount of bandwidth and will lead to slow links with large networks.
- RIPv1 and RIPv2 converge very slowly. It can take many minutes, which may cause inconsistencies in routing. This is unsuitable in large network environments.
- RIPv1 and RIPv2 do not have concept of network delays and link costs. It bases its routing decisions on hop counts. The path with the lowest hop count to the destination is always preferred even when the longer path has a better aggregate link bandwidth and lower delays.
- RIPv1 and RIPv2 are flat networks without areas or boundaries.

OSPF is a link-state protocol. A router using link-state protocol determines the state or status of its link then, constructs a link state advertisement (LSA) with the status of its links and transmits this information to its neighbor. Using this method, routers are able to build a complete list of all the routes to all destinations by compiling LSAs from each router. The OSPF routers will use this information to construct a table so that each router can identify the location of each subnet, the routers that are directly connected to it, and the path of a specific router.

Deleted: . Each router

OSPF is also an interior gateway protocol that distributes routing information between routers in a single autonomous system. After all the routers have formed their databases based on the LSA information, they will run the shortest path first

algorithm. A tree structure is formed and the shortest paths to all other destinations are mapped out. The selection of the path to these destinations is based on metrics. Metrics may be based on hop count, bandwidth, load, cost, reliability, delay, or controlled statically by the user. It can provide the network managers and administrators with greater control over the routing path in the network.

OSPF provides several networking features that enable a more robust and flexible internetworking environment. The features are:

- No hop counts limitation.
- IP multicast to send link-state updates. This ensures less processing on routers that are not listening for OSPF packets.
- "Event triggered" updates. These are sent only when there are routing changes in the network.
- Better load balancing.
- Logical definition of networks in a hierarchical network structure where routers can be divided into areas. This will cut down the unnecessary propagation of subnet information.

**OSPF Packet Format**

All OSPF packets begin with a 24-byte header. OSPF packets consist of nine fields.

Field Length in Bytes									
1	1	2	4	4	2	2	8	Variable	
Ver Num	Type	Packet Length	Router ID	Area ID	Check-sum	Authen-tication Type	Authentication	Data	

**Figure 8.** OSPF Packets

The following is a description of the OSPF header fields:

- **Version Number**  
This identifies the OSPF version.
- **Type**  
This identifies the OSPF packet type which includes the following:
  - **Hello**  
Used to establish and maintain the neighbor relationships.

- **Database Description**  
Used to describe the topological database contents. When an adjacency is initialized, these messages are exchanged.
  
- **Link-State Request**  
Used to request for the topological database from the neighboring routers. When router discovers any part of its topological database is outdated, these messages are exchanged.
  
- **Link-State Update**  
Used to respond to a link-state request packet, it is also used for the regular dispersal of LSAs. A single link-state update packet can contain several LSAs.
  
- **Link-State Acknowledgement**  
Used to acknowledge the link-state update packets.
  
  
- **Packet Length**  
This specifies the packet length that includes the OSPF header in bytes.
  
  
- **Router ID**  
This identifies the source of the packet.
  
  
- **Area ID**  
This identifies the area to which the packet belongs. All OSPF packets are associated with a single area.
  
- **Checksum**  
This checks the entire packet contents for errors and damage in transit.
  
  
- **Authentication Type**  
This contains the authentication type. All OSPF protocol exchanges are authenticated. The authentication type is configurable on per-area basis.
  
  
- **Authentication**  
This contains authentication information.
  
  
- **Data**  
This contains encapsulated upper-layer information.

### **OSPF Operation**

1. The Shortest Path First (SPF) routing algorithm is the basis for OSPF operation. When a SPF router is powered up, it will initialize its routing protocol data structures and wait for indications from the lower-layer protocols to confirm its interfaces are functional.
2. After a router is assured that its interfaces are functioning, it uses the OSPF Hello protocol to acquire neighbors, which are routers with interfaces to a common network. The router sends and receives hello packets to and from its neighbors. Hello packets allow the routers to know that other routers are still functional.
3. For networks with more than two routers, the Hello protocol will elect a designated and backup designated router. The designated router is responsible for generating LSAs for the entire multi-access network. Designated routers allow a reduction in the size of the topological database and in network traffic.
4. Each router periodically sends an LSA to provide information on a router's adjacencies or to inform others when a router's state changes. By comparing established adjacencies to link states, failed routers can be detected quickly and the network's topology can be altered appropriately.
5. From the topological database generated from LSAs, each router calculates a shortest path tree using itself as root. The shortest path tree, in turn, yields a routing table.

OSPF also supports equal-cost, multipath routing, and routing based on upper-layer type-of-service (TOS) requests. For example, if an application specifies that certain data is urgent, the OSPF will transport this urgent datagram if it has high-priority links at its disposal.

### **Advantages and Disadvantages of OSPF**

OSPF can provide fast updates to all routers on large and complex networks. It is able to forward traffic across multiple paths to a single destination. Unlike RIP which depends on hop count, OSPF uses metrics to determine the best path to a destination.

With the ability to handle large and complex networks, network managers and administrators will be faced with greater complexity in configuring and setting up OSPF networks. Greater computing power of routers will also be required. Network managers and administrators must understand, manually calculate, and factor in path costs that the routers will use during router-to-router exchanges.

**Summary: Open Shortest Path First**

OSPF is a powerful routing protocol and is rich in features because of its flexibility. It provides high functionality open protocol standard that enables inter-vendor networking with TCP/IP protocol suite. OSPF has faster convergence compared to RIPv1 and RIPv2. It can handle large and complex networks, and supports authentication, hierarchical segmentation, route summarization and aggregation. With Layer 3 switches, routing performance can be significantly enhanced, thus reducing the dependency on computing power in routers.

## **Virtual Router Redundancy Protocol**

Virtual Router Redundancy Protocol (VRRP) is a routing protocol developed to provide redundancy for the routers on a LAN. It is designed to eliminate the single point of failure inherent in static routing environments by providing alternative routing paths automatically.

In many networks, a common practice is to have one router to serve as the router for forwarding packets from a group of hosts on a LAN. This practice, however, does not provide redundancy and if the router fails, there is no way to immediately use another router as a backup.

Using VRRP, a virtual IP address can be specified manually or dynamically using Dynamic Host Configuration Protocol (DHCP) to share among routers, with one designated as the Master and the others as backups. When the Master fails, an election process takes place and one of the backups will be selected for dynamic failover. The virtual IP address will then be mapped to this backup router's IP address, and this backup router will become the Master router taking over the forwarding responsibility.

Using this approach, the host's gateway information need not be changed regardless of which path is used. This will mean administrative overheads can be significantly reduced in comparison with redundancy schemes that require hosts to be configured with multiple default gateways.

### **VRRP Terms and Definitions**

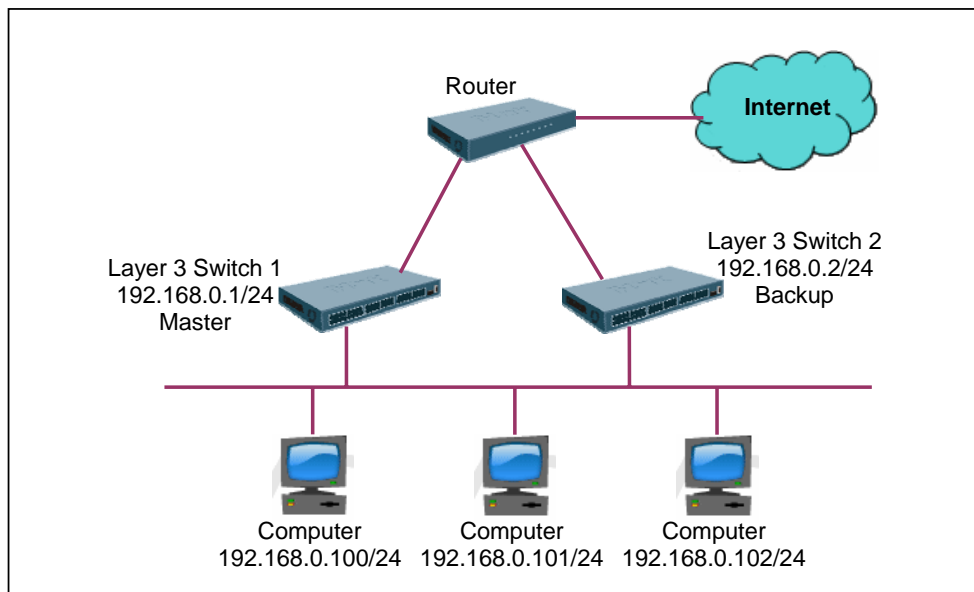
Before looking into examples of how VRRP works, a definition of VRRP terms must first be explained. The following is a summary of VRRP definitions:

Component	Definition
Virtual Router	An IP address or set of IP addresses shared by VRRP Routers and addressed by a user configured unique identifier.
VRRP Router	A router running Virtual Router Redundancy Protocol.
IP Address Owner	The VRRP router that has its virtual IP address configured as its real interface address. Only one IP Address Owner is allowed in a Virtual Router. The IP Address Owner always wins the election process to become Master in a Virtual Router. IP Address Owner is an optional designation.

Renter	VRRP Routers that are not configured with the Virtual Router's IP address as their real interface address.
Master	The Master is the router that assumes responsibility for forwarding packets sent to the Virtual Router and answering address resolution requests for the Virtual Router.
Backup	This is the available VRRP Router that assumes the Master duties should it fails.

### VRRP Operation

The example below illustrates how VRRP works using a simple two-node Virtual Router.



**Figure 9.** Illustration on VRRP

The routers are configured as VRRP Routers, and together they form a Virtual Router. Layer 3 switch 1 has its real interface configured with the IP address of the Virtual Router 192.168.0.1/24, and it is the IP Address Owner. As the IP Address Owner, the VRRP Router will be the Virtual Router Master so long it is available and active.

Layer 3 switch 2 is the Virtual Router Backup. Its real interface is configured with an IP address of the same subnet as the Virtual Router, but not to the Virtual Router IP address. It is a Renter and also a Virtual Router Backup.

The Virtual Router is assigned a Virtual Router ID (VRID). If the VRID is 1, both VRRP Routers will use the MAC address of 00-00-5E-00-01-01.

The default gateways of the computers are configured with the Virtual Router's IP address. The Master will forward the packets to the destination subnets and responds to the Address Resolution Protocol (ARP) request. As the Master is also the Virtual Interface Router's IP Address Owner, it will also respond to the Internet Control Message Protocol (ICMP) ping requests and IP datagrams destined for the Virtual Interface Router's IP address. The Backup will not forward any traffic or respond to ARP requests on behalf of the Virtual Interface Router.

When the Master is not available, the Backup becomes the Master and takes over its responsibilities. It will perform ~~packet~~ forwarding and respond to the ARP requests. However, the new Master router is not configured as the IP Address Owner, so it will not respond to the ICMP ping requests or IP datagrams destined to that address.

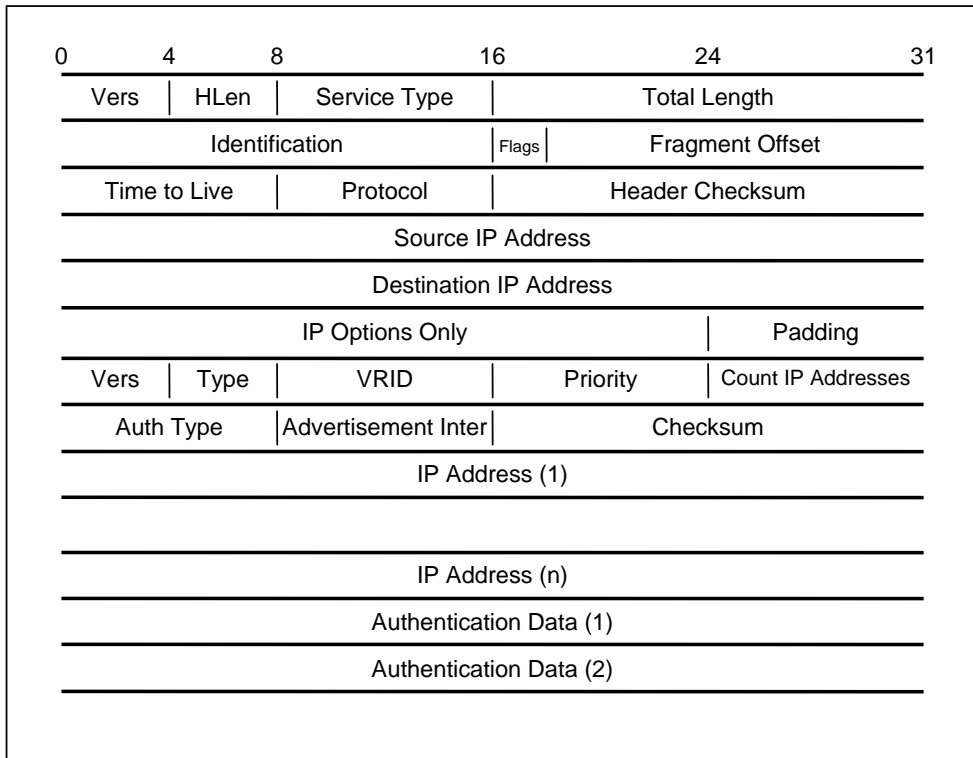
Deleted: ort

Each VRRP Router that is a Renter is configured with a priority between 1 to 254. The Owner has a priority of 255. In the above example, this will mean that Layer 3 Switch 1 owns the Virtual Router's IP address and will declare itself as the Master and sends out an advertisement to all other VRRP Routers. The IP Address Owner is always the Master as long as it is available because when the original Master fails, the election process will always select the IP Address Owner by default.

There is no necessity for the Virtual Router IP address to be owned by a VRRP Router connecting the LAN to the Internet in the example above. The bidding process that determines the Master in this case will be different. The process will involve comparing two criteria. First, the VRRP Router with the highest priority will become the Master. Second, if the priorities are the same, the highest IP address will be the Master. To understand the bidding process, it is essential that VRRP advertisement packet must first be understood.

### **VRRP Packets**

VRRP packets are sent encapsulated in IP packets. The figure below shows the layout of the packet's IP header and the packet itself.



**Figure 10.** Contents of a VRRP Packet

The important fields in the VRRP IP header are:

- **Source IP Address**

This 32-bit source address is the primary IP address of the interface from which the packet is being sent. This is the IP address of the Master router's interface connected to the LAN.

- **Destination IP Address**

This is a 32-bit IP multicast address 224.0.0.18 that is assigned by the Internet Assigned Numbers Authority (IANA) for VRRP. All the routers running VRRP receive this multicast.

- **Time to Live (TTL)**

This 8-bit field has a value equal to 255. VRRP packet received with TTL not equal to 255 will be discarded.

- **Protocol**

This is an 8-bit protocol field. The IP protocol number assigned by IANA for VRRP is 112.

- **Version**

This is a 4-bit VRRP version field. The available version is 2.

- **Type**

This is a 4-bit VRRP packet field. The only type available is ADVERTISEMENT.

- **Virtual Router Identifier (VRID)**

VRID is used to identify the virtual router for which the packet is reporting status.

- **Priority**

This 8-bit field specifies the sending VRRP router's priority for the Virtual Router. The priority value of the VRRP router that owns the IP address associated with the virtual router must be 255. The default priority value is 100. The assigned values can be between 1 to 254, where the higher value means higher priority. A priority of 0 means the Master has stopped functioning, and the Backup Router needs to transit to the Master state.

- **Count IP Addresses**

This 8-bit field specifies the number of IP addresses contained in this VRRP advertisement.

- **Authentication Type**

This 8-bit field specifies the authentication type being used. Any packet that has an unknown authentication type or does not match the locally configured authentication type will be discarded.

There are three authentication methods defined. The first is No Authentication with the value 0. This means the VRRP protocol exchanges are not authenticated. It is set to 0 on transmission and ignored on reception.

The second is Simple Text Password authentication. The VRRP protocol exchanges are authenticated by a clear text password. The content in the Authentication Data field is set to a locally configured password on transmission. The receiving VRRP Router must check that the Authentication Data in the packet

and match with its configured authentication string. Packets that do not match are discarded.

The third is IP Authentication Header authentication. This means the VRRP protocol exchanges are authenticated using the method defined by the IP Authentication Header.

- **Advertisement Interval**

This 8-bit field specifies the time interval between advertisements sent from the master to let the backup router know that it is alive. All routers with the same VRID should have the same advertisement interval.

- **Checksum**

This 16-bit checksum is used to detect data corruption in the VRRP message.

- **IP Address**

This is the Virtual Router's IP address that the Master is backing up. Each address associated with the Virtual Router is included in a separate 32-bit field within the announcement.

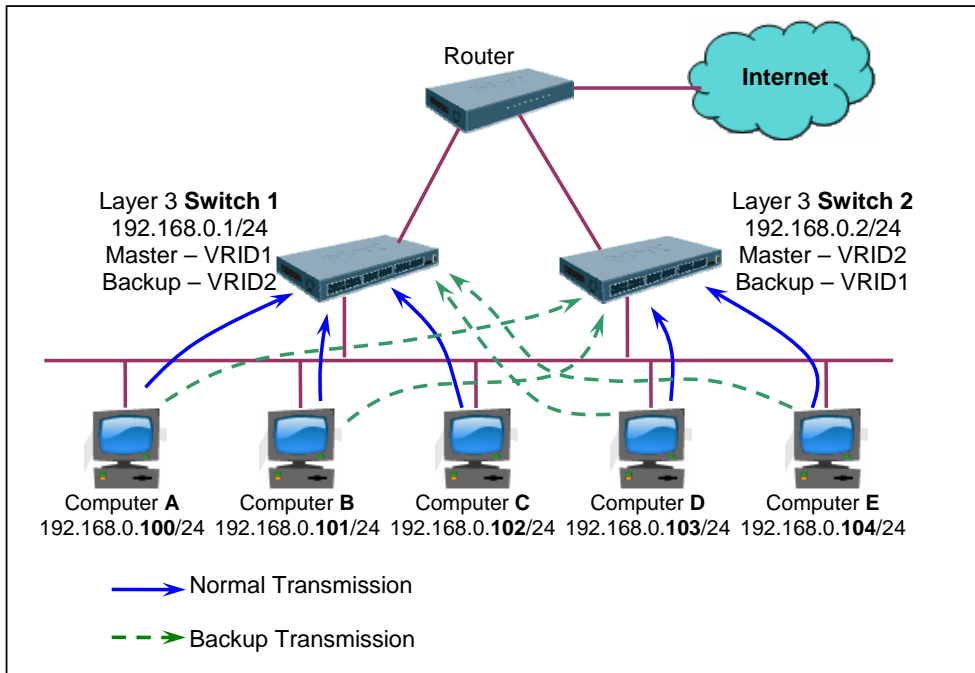
- **Authentication Data**

This is currently used for Simple Text Authentication. It can input up to 8 characters of plain text. It must match the locally configured string or it will be discarded.

### **Applications for VRRP**

In our previous example, the Master router is performing all the work, while the other router is an idle Backup. In order to utilize the bandwidth efficiently, two different VRIDs can be created to provide sharing of load.

In the example below, part of the traffic can be sent through Layer 3 Switch 1 and other traffic through Layer 3 Switch 2 to perform load balancing.



**Figure 11.** Load Sharing with VRRP

The Layer 3 Switch 1 is the default gateway for Computer A, B, and C and Layer 3 Switch 2 is the default gateway for Computer D and E. There are two VRIDs: VRID 1 and VRID 2. Switch 1, with VRID 1, is the Master for Computer A, B, and C and Backup for Computer D and E. Switch 2, with VRID 2, is the Master for Computer D and E and Backup for Computer A, B, and C. In this example, the traffic going out of the LAN 192.168.0.0/24 subnet is shared between the two Layer 3 switches, thus efficiently utilizing the routers and bandwidth.

VRRP can also be used to eliminate downtime due to maintenance, such as software upgrades, which will need rebooting. Depending on the complexity of router configuration, the maintenance downtime or the Mean Time To Restoration (MTTR) can range from 5 to 40 minutes. During this period, the router will be unable perform its regular functions. Using VRRP solves this problem. In the example above, if Switch 1 is down for maintenance, Switch 2 can take over the routing function temporarily.

**Summary: Virtual Redundancy Routing Protocol**

VRRP is a protocol that allows dynamic failover of a router to a Backup when the Master router is down. Its primary function is to provide redundancy routing using a common virtual IP address shared by several routers. When the Master router fails, VRRP goes through an election process to appoint a Backup router to take over the role of the Master automatically. Using VRRP, routers can perform load sharing, and eliminate downtime due to router maintenance.