Redundancy and load balancing at L3 in Local Area Networks

Fulvio Risso

Politecnico di Torino
Default gateway redundancy (1)
Default gateway redundancy (2)

- Objective
  - Provide redundancy of the default gateway
  - Two routers on the LAN are useless if hosts are not able to switch to another one in case the first fails
    - Hosts on the LAN are not able to learn the network topology through Layer 3 routing protocols

- Please note that incoming traffic (from the Internet to the LAN) will arrive correctly, even if one router fails
  - We are having troubles only in the opposite direction
Default gateway redundancy (3)

- Solutions
  - HSRP (Hot Standby Routing Protocol)
    - Cisco proprietary protocol defined in RFC 2281
  - VRRP (Virtual Router Redundancy Protocol)
    - Standard protocol defined in RFC 3768
  - GLBP (Gateway Load Balancing Protocol)
    - Cisco proprietary protocol
HSRP: overview

Objectives

- First, redundancy of the default gateway
- Second, load balancing
HSRP: the idea

ARP Request: who has IP 2.2.2.4?
ARP Reply: IP 2.2.2.4 is at MAC 00:00:00:44:44:44

"Hello" packets

R3 (active)
IP: 2.2.2.3
MAC: 00:00:00:33:33:33
HSRP Group: 4
Virtual IP: 2.2.2.4
Virtual MAC: 00:00:00:44:44:44

R2 (stand-by)
IP: 2.2.2.2
MAC: 00:00:02:22:22:22

R1 (listen)
IP: 2.2.2.1
MAC: 00:00:00:11:11:11
HSRP Group: 4

Internet
HSRP: working operations (1)

- Routers that provide redundancy belong to the same “group”
  - Each group emulates a single virtual router

- Routers can be
  - Active: the one that has the right to serve the LAN
  - Stand-by: the one that takes place in case the Active fails
  - Listen: the others that are neither Active nor Stand-by
    - May become stand-by in case the active fails

- Within each group, we define two additional (shared) addresses
  - A virtual IP address
  - A virtual MAC address
HSRP: working operations (2)

The interface of the Active router has the followings assigned addresses:

- Primary IP address (inserted in the source IP header field)
- Virtual IP address (used by hosts as default gateway)
- Physical MAC address assigned by the NIC manufacturer
- Virtual MAC address allocated to the HSRP Group
  - Well-known MAC address, derived by the HSRP group

The interfaces of the Standby and Listen routers have the followings assigned addresses:

- Primary IP address (inserted in the source IP header field)
- Physical MAC address assigned by the NIC manufacturer
HSRP: working operations (3)

- Only the Active Router can use the virtual addresses
  - Only the Active router can reply to the virtual IP/MAC addresses
  - It does no longer use its IP/MAC addresses (unless an explicit request for that address is received)

- Standby and Listen routers reply only to their own IP/MAC addresses
  - The Standby will start using the virtual MAC when it is promoted to “active”
HSRP: working operations (4)

- Keep-alive message
  - Used to elect the router current in “active” mode
    - Priority
    - (or, if equal) Higher IP address
HSRP: Virtual MAC address

- Taken in charge by the Active router and used to answer at the received ARP request packets
- Unicast address
- Token Ring has only 3 possible values (corresponding to HSRP Group 0, 1, 2)
  - C0-00-00-01-00-00
  - C0-00-00-02-00-00
  - C0-00-00-04-00-00
- Well known virtual MAC address for other LANs (e.g., 802.3, 802.11 etc.)
  - 00-00-0C-07-AC-xx
    - xx represents the HSRP Group
HSRP packet format

```
| MAC header | IP header | UDP header | HSRP | FCS |
```

<table>
<thead>
<tr>
<th>0</th>
<th>8</th>
<th>16</th>
<th>24</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>Op Code</td>
<td>State</td>
<td>Hello-time</td>
<td></td>
</tr>
<tr>
<td>Hold-time</td>
<td>Priority</td>
<td>Group</td>
<td>Reserved</td>
<td></td>
</tr>
</tbody>
</table>

* Authentication Data
* Authentication Data
* Virtual IP Address
HSRP Encapsulation

- HSRP packets
  - Required for the “keepalive” function between routers
  - Encapsulated in UDP, src/dst port is 1985
- IP Layer
  - Transmitted to multicast address 224.0.0.2 (“all routers”)
  - Source IP address is the actual IP address of the router (in order to elect the best router as active)
  - TTL = 1 (packets are not further forwarded by routers)
- MAC layer
  - Destination address: derived from the Destination IP address
  - Source MAC: Virtual MAC address for the group
HSRP header: “opcode” field

- Describes the type of message contained in this packet
- Possible values
  - 0 = Hello
    - The router is running and is capable of becoming the active or standby router
  - 1 = Coup
    - The router wants to become the active router
  - 2 = Resign
    - The router does no longer want to be the active router
    - “Coup” and “Resign” are not necessarily used together
      - E.g. a router that has highest priority can send the “coup” message to take over the current router, but this router goes into “speak” state without sending any “resign”
HSRP header: “state” field (1)

- Describes the current state of the router sending the message

- Possible values
  - 0 = Initial
    - This is the starting state and indicates that HSRP is not running
  - 1 = Learn
    - The router has not determined the virtual IP address and is still waiting to hear it from the active router
  - 2 = Listen
    - The router knows the virtual IP address, but is neither the active router nor the standby router
HSRP header: “state” field (2)

- Possible values (cont.):
  - 4 = Speak
    - The router sends periodic Hello messages and is actively participating in the election of the active and/or standby router
  - 8 = Standby
    - The router is a candidate to become the next active router and sends periodic Hello messages
    - At most one router can be in Standby state (for each group)
  - 16 = Active
    - The router is currently forwarding packets that are sent to the group virtual MAC address
    - Each group has at least one (and only one) router in Active state
HSRP header: “Hello time” and “Hold time”

- **Hello-Time**
  - Period between the Hello messages sent by the routers
  - In case no active routers exists, a router uses its configured value (default: 3 seconds)

- **Hold-Time**
  - Validity of the current Hello message
  - When this timer expires, the Standby router proposes itself as Active router
  - In case no active routers exists, a router uses its configured value (default: 10 seconds)

- **Notes**
  - These numbers are set by the active router for the entire LAN
  - The convergence time for an HSRP network is about 10 seconds
HSRP header: “priority” and “group” fields

- **Priority**
  - Used to force the election of the router with highest priority (higher number means higher priority)
  - In the case of routers with equal priority, the router with the highest IP address wins
  - Default Priority is 100

- **Group**
  - Group ID the current HSRP instance is referring to
  - For Token Ring, values between 0 and 2 are valid
  - For other media, values between 0 and 255 are valid
    - Max 255 groups
HSRP header: Authentication and Virtual IP Address

- Authentication Data
  - This field contains a clear-text 8-characters password
  - If no authentication data is configured the default text is “cisco”
  - Really simple authentication; mostly used to differentiate multiple instances of HSRP within the same LAN

- Virtual IP address
  - The virtual IP address used by this group
    - IP address of the default gateway configured on the hosts
  - If the virtual IP address is not configured on a router, then it may be learned from the Hello message sent from the active router
HSRP packets over the LAN

R2 (active)
IP: 2.2.2.2
MAC: 00:00:00:22:22:22
HSRP Group: 1
V-IP: 2.2.2.3
V-MAC: 00:00:0C:07:AC:01

H1
DG: 2.2.2.3

<table>
<thead>
<tr>
<th>ETH.SRC: 00:00:0C:07:AC:01</th>
<th>ETH.DST: 01:00:5E:00:00:02</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP.SRC: 2.2.2.2</td>
<td>IP.DST: 224.0.0.2</td>
</tr>
<tr>
<td>UDP.SPORT: 1985</td>
<td>UDP.DPORT: 1985</td>
</tr>
<tr>
<td>HSRP HELLO (HSRP Group 1)</td>
<td></td>
</tr>
</tbody>
</table>

R1 (standby)
IP: 2.2.2.1
MAC: 00:00:00:11:11:11
HSRP Group: 1

Multicast MAC address
(derived from 224.0.0.2)
HSRP: fault reaction

Data to Internet
IP.dst: IP_remote
MAC.dst: 00:00:0C:07:AC:01

"Hello" packets

Internet

R1 (standby)
IP: 2.2.2.1
MAC: 00:00:00:11:11:11
HSRP Group: 1
Virtual IP: 2.2.2.3
Virtual MAC: 00:00:0C:07:AC:01

R2 (active)
IP: 2.2.2.2
MAC: 00:00:00:22:22:22
HSRP Group: 1
Virtual IP: 2.2.2.3
Virtual MAC: 00:00:0C:07:AC:01

H1
DG: 2.2.2.3

H2
DG: 2.2.2.3

H3
DG: 2.2.2.3

Virtual IP: 2.2.2.3
Virtual MAC: 00:00:0C:07:AC:01
HSRP and interfaces (1)

- HSRP has a per-interface configuration
  - Interfaces belonging to the same group must be reachable at L2
  - HSRP works also if the two interfaces belong to the same device
HSRP and interfaces (2)

- HSRP can be used to implement Dual homing servers/routers and achieve single-machine redundancy
  - Two standard NICs
  - HSRP will handle the virtual IP/MAC address on the primary interface
  - Both interfaces must belong to the same LAN

- Other option for redundancy
  - Server (fault tolerant) NIC
    - Not available on routers
HSRP and IP networks

- HSRP is specific for a given IP network
  - A LAN with two Logical IP Networks must use two HSRP groups

![Diagram of HSRP and IP networks]

**MAC: 00:00:00:11:11:11**
**IP: 1.1.1.1**
**HSRP Group: 1**
**IP: 3:3:3:3**
**HSRP Group: 4**

**MAC: 00:00:00:22:22:22**
**IP: 1.1.1.2**
**HSRP Group: 1**
**V-IP: 1.1.1.254**
**V-MAC: 00:00:0C:07:AC:01**
**IP: 3:3:3:4**
**HSRP Group: 4**
**V-IP: 3.3.3.254**
**V-MAC: 00:00:0C:07:AC:04**
HSRP and VLANs

- Each VLAN is a separated LAN, with its own default gateway
- Multiple HSRP groups are required
HSRP and Filtering database

- Only the active router can use the Virtual MAC as source address in Hello Messages
- When the Active router changes, the new Active router will start sending messages with that MAC address
- All the switches will update their filtering database accordingly

<table>
<thead>
<tr>
<th>MAC</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>07:AC:01</td>
<td>Fe0</td>
</tr>
<tr>
<td></td>
<td>Fe1</td>
</tr>
</tbody>
</table>

R1 (standby)
- IP: 2.2.2.1
- MAC: 00:00:00:11:11:11
- HSRP Group: 1
- Virtual IP: 2.2.2.3
- Virtual MAC: 00:00:0C:07:AC:01

R2 (active)
- IP: 2.2.2.2
- MAC: 00:00:00:22:22:22
- HSRP Group: 1
- Virtual IP: 2.2.2.3
- Virtual MAC: 00:00:0C:07:AC:01

Internet
**HSRP and ARP caches**

- A gratuitous ARP Reply (in broadcast) is sent when a router becomes Active
  - Source MAC: Virtual MAC address
  - Destination MAC: broadcast
  - ARP Reply: Virtual IP is at Virtual MAC

- Strictly speaking, not needed
  - ARP mapping do not change if the Virtual IP – Virtual MAC are used
  - In some (old) cases, we may wanted to use the MAC address of the router instead of the virtual IP
  - In those cases, we need to update the ARP cache
    - Although some Operating Systems may not handle it properly (because of the unexpected broadcast MAC address)
HSRP and network utilization

- Only one exit link used
  - Unused bandwidth on the stand-by router
- Asymmetric routing
  - The external routing protocol can use both paths
HSRP and load sharing (1)
HSRP and load sharing (2)

- Multiple HSRP groups
  - Multi-group HSRP (mHSRP)
  - One router active for the first group, the other active for the second group
  - Clients are configured half with the first DG, half with the second

- Note that load sharing in ingress does not depend on HSRP
  - Ingress traffic is handled by the routing protocol on the WAN side (e.g. BGP multi-homing)

- Is it really needed?
  - Often corporate networks have much more incoming than outgoing traffic
  - In this case, the egress bandwidth of a single link may be enough
HSRP and load sharing (3)

- Problems of the achieved load balancing
  - Load balancing is statically defined by physically partitioning hosts between two different default gateways
    - What about if one group generates much more traffic than the other?
  - Configuration burden
    - Not easy to differentiate the Default Gateway on the clients
    - Usually DHCP is used, and this usually returns a single DG for all hosts

- Asymmetric routing in any case
  - Ingress and egress paths may be different, with or without mHSRP
  - Depends on the external routing protocol (OSPF, BGP, etc.)
Load sharing with VLANs

- We can achieve the same results of mHSRP
- However, no protection in case the router fails
HSRP: Preemption capability

- Parameter configured on each router
- If a router has higher priority than the active router and preemption is configured, it MAY take over as the active router using a Coup message
  - Without preemption, the currently Active router will stay active until it has a fault
  - A router configured with the highest priority cannot force the current Active router to resign, unless preemption is used
HSRP: basic configuration

R2(config)# interface ethernet 0
R2(config-if)# ip address 10.1.1.2 255.255.255.0
R2(config-if)# standby 24 ip 10.1.1.5
R2(config-if)# standby 24 preempt

R2(config)# interface ethernet 0
R2(config-if)# ip address 10.1.1.2 255.255.255.0
R2(config-if)# standby 24 ip 10.1.1.5
R2(config-if)# standby 24 preempt

R1(config)# interface ethernet 0
R1(config-if)# ip address 10.1.1.1 255.255.255.0
R1(config-if)# standby 24 ip 10.1.1.5
R1(config-if)# standby 24 preempt

HSRP is not a router-wide function; it is a propriety of the interface that must be on the LAN.
HSRP: “track” function (1)

- Problem: a failure on the WAN link does not trigger the Stand-by router to take place
  - Packets are sent to R1, from there to R2 (routing protocol know the possible routes)
    - Opposite route works as well (not under HSRP control, though)
  - No problems in connectivity, but additional overhead in forwarding
HSRP: “track” function (2)

- “Track” the status of the physical layer on the interface
  - Dynamically decrease the HSRP Priority when a tracked interface goes down
    - By default HSRP algorithm decrease the Priority by 10 when the link-layer of a tracked interface goes down

- Be careful: only some types of faults cause the interface to go down
  - E.g. an interface will stay up if connected to an active L2 switch
HSRP: advanced configuration

R1(config)# interface ethernet 0
R1(config-if)# ip address 10.1.1.1 255.255.255.0
R1(config-if)# standby 1 ip 10.1.1.5
R1(config-if)# standby 1 priority 105
R1(config-if)# standby 1 preempt
R1(config-if)# standby 1 track Serial0
R1(config-if)# standby 2 ip 10.1.1.6
R1(config-if)# standby 2 preempt
R1(config-if)# standby 2 track Serial0

R2(config)# interface ethernet 0
R2(config-if)# ip address 10.1.1.2 255.255.255.0
R2(config-if)# standby 1 ip 10.1.1.5
R2(config-if)# standby 1 preempt
R2(config-if)# standby 1 track Serial0
R2(config-if)# standby 2 ip 10.1.1.6
R2(config-if)# standby 2 priority 105
R2(config-if)# standby 2 preempt
R2(config-if)# standby 2 track Serial0
VRRP overview (1)

- A “smart” clone of HSRP, with care taken not to infringe any Cisco patent
- Functioning, philosophy is exactly the same as HSRP
  - Same way to achieve the Default Gateway redundancy
  - Same way to achieve Load Balancing
VRRP overview (2)

- Minor differences
  - Packet encapsulated in IP, protocol type 112 (no longer in UDP)
  - Transmitted to multicast address 224.0.0.18
  - Different MAC addresses associated to each group
  - TTL = 255
    - A VRRP router receiving a packet with the TTL not equal to 255 must discard the packet (only one possible hop)
  - Active/Standby → Master/Backup
  - Hello Messages → Advertisement messages
  - HSRP Group → Virtual Router ID (VRID)
  - Some timers (see later)
VRRP overview (3)

- Major (?) differences
  - Each master VRRP router can control more than one IP Address
  - A VRRP Router may backup one or more virtual routers
  - Any of the virtual router's IP addresses on a LAN can then be used as the Default Gateway by end-hosts
  - Support multiple logical IP subnets on a single LAN segment
  - For any VRID a single Master Router is elected the remaining routers are selected as Backup Routers (no longer routers in “Listen”)
  - Only the Master router sends Advertisement packets
  - The master router may have the same address as the group virtual router address
  - “tracking” not available
  - “preempt” is the specified behavior
VRRP: packet format (1)
VRRP: packet format (2)

- IP header
  - Source IP: real IP address of the interface the packet is being sent from
  - Destination IP: 224.0.0.18

- Type
  - The type field specifies the type of this VRRP packet. The only packet type is:
    - 1 Advertisement

- VRID
  - The Virtual Router Identifier (VRID) field identifies the virtual router this packet is reporting status for
    - Allowed values: 1-255
**VRRP: packet format (3)**

- **Priority**
  - Router with highest priority will become the master
    - In case of a tie, the router with the highest real IP becomes master
  - Priority = 255: assigned automatically to the router that has the same address as the virtual router
    - The router will be the master router (known as the “virtual address owner”)
  - Priority = 1-254: normal priority values
  - Priority = 0 → the current router does not participate in VRRP
    - Also advertised during an orderly shutdown of a master in order to speed-up Backup promotion (no need to timeout)
  - The default priority value for VRRP routers backing up a virtual router is 100
VRRP: packet format (4)

- Count IP Addr
  - number of IP addresses contained in this VRRP Advertisement

- Authentication Type
  - Unused in the current version (RFC 3768)
  - Removed because operational experience showed that they did not provide any real security and would only cause multiple masters to be created
  - 0 = No Authentication
VRRP Timers

- Advertisement Interval
  - Time interval (in seconds) between Advertisements
  - default value = 1 s (HSRP was 3)

- Skew_Time
  - \((256 - \text{Priority}) / 256\) (seconds)

- Master_Down_Interval
  - \((3 \times \text{Advertisement Interval}) + \text{Skew time}\)
  - Time interval for Backup to declare Master down (seconds)
  - After that time, a new Master is elected
  - In case an orderly shutdown of a master is detected, backup waits only for the skew time
VRRP: Virtual MAC Address

- Well known virtual MAC address for any LAN except Token Ring (e.g. 802.3, 802.11 etc.)
  - 00-00-5E-00-01-xx
    - xx represents the VRID
  - OUI changed (C0-00-00 is owned by Cisco)
VRRP: Virtual MAC Address for Token Ring

<table>
<thead>
<tr>
<th>VRID</th>
<th>Token Ring Functional Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>03-00-02-00-00-00</td>
</tr>
<tr>
<td>2</td>
<td>03-00-04-00-00-00</td>
</tr>
<tr>
<td>3</td>
<td>03-00-08-00-00-00</td>
</tr>
<tr>
<td>4</td>
<td>03-00-10-00-00-00</td>
</tr>
<tr>
<td>5</td>
<td>03-00-20-00-00-00</td>
</tr>
<tr>
<td>6</td>
<td>03-00-40-00-00-00</td>
</tr>
<tr>
<td>7</td>
<td>03-00-80-00-00-00</td>
</tr>
<tr>
<td>8</td>
<td>03-00-00-01-00-00</td>
</tr>
<tr>
<td>9</td>
<td>03-00-00-02-00-00</td>
</tr>
<tr>
<td>10</td>
<td>03-00-00-04-00-00</td>
</tr>
<tr>
<td>11</td>
<td>03-00-00-08-00-00</td>
</tr>
</tbody>
</table>
HSRP/VRRP: convergence

- HSRP: about 10 sec with default parameters
  - Hold Time
  - User can configure Hello-Time and Hold-Time to improve this value

- VRRP: about 4 sec with default parameters
  - \((3 \times \text{Advertisement\_Interval}) + \text{Skew\_time}\)
  - User can configure Advertisement Interval and Priority
  - Less flexible than HSRP
HSRP/VRRP on real LANs: L2 resiliency (1)

Fault 1: OK (traffic re-routed through R2)
Fault 2: OK (traffic re-routed through R2)
Fault 3: OK (traffic continues through R1)
Fault 4: non-optimized exit path, but everything works ("track" function useless)
Fault 5: IP network partitioned in half (exit traffic may use R1, but incoming traffic may come through R2)

We must engineer the network with resiliency at data-link level (e.g. link aggregation between SW-1 and SW-2 or STP/RSTP)
HSRP/VRRP on real LANs: L2 resiliency (2)

- HSRP/VRRP do not protect from all faults on the L2 network
- Solutions: use STP/RSTP or link aggregation
  - The latter is the best for its reduced convergence time
  - STP may take 50s to converge; during this period, malfunctioning may occur
    - IP networks still partitioned
- HSRP/VRRP do not protect from some faults on the WAN link
HSRP/VRRP on real LANs: flooding (1)

- Switches may have incomplete filtering database
  - Some entries may be missing, due to the Aging Time
  - If a frame directed to that host is received, the frame is flooded
- A periodic generation of broadcast frames avoids the problem
  - Some hosts generate a limited number of broadcasts (e.g. UNIX or VmWare)
- Sometimes ARP messages are sent only “occasionally” on the network
  - ARP cache is often 5 minutes or more
  - Max Ageing Time in the filtering database is often 2 minutes
HSRP/VRRP on real LANs: flooding (2)

- A possible pathological situation
  - An host A that wants to contact the station B has still the mapping \((IP(B), MAC(B))\) in its ARP cache
  - The entry related to MAC(B) is no longer in the filtering database
  - In that case, host A will not send an ARP Request to B, and the MAC frame is generated and sent on the network
    - Flooding
  - Please note that in any case, the ARP Reply may not reach the entire network, although usually this is not a problem
- This situation may be extremely common with HSRP/VRRP
HSRP/VRRP on real LANs: flooding (3)

Switches on this path will refresh the entries in the filtering database at each received frame (e.g. unicast)

Switches on this path will refresh the entries in the filtering database only when host H1 generates a broadcast frame
HSRP/VRRP on real LANs: flooding (4)

- Hypothesis
  - Let us suppose that host H1 does not generate periodic broadcasts on its own
  - Egress and ingress routers are different
    - Asymmetric routing
  - The ARP Cache on router R2 last longer than the filtering database

- Description
  - Egress traffic from station H1 updates the filtering database on the exit path
  - The ingress router still has a valid mapping for H1 in its ARP cache
  - Switches on the ingress path do no longer have the H1 entry in the filtering database
  - The only option (for the switch) is to send the frame on all its ports
  - \(\rightarrow\) periodic flooding (every now and then)
HSRP/VRRP on real LANs: flooding (5)

- Solutions
  - Re-engineer the L2 spanning tree (not really a solution)
  - Force stations to send broadcast frames rather often (< Max Ageing Time)
  - Increase Max Ageing Time on the switches
HSRP/VRRP on unidirectional links (1)

R1 (standby ↞ active)
IP: 2.2.2.1
MAC: 00:00:00:11:11:11
→ V-IP: 2.2.2.3
→ V-MAC: 00:00:0C:07:AC:01

R2 (active)
IP: 2.2.2.2
MAC 00:00:00:22:22:22
V-IP: 2.2.2.3
V-MAC: 00:00:0C:07:AC:01

Instability in the filtering database of the switch
HSRP/VRRP on unidirectional links (1)

- R1 will not receive HSRP packet from R2, therefore both will become active
  - Both R1 and R2 will send HSRP Hello using the virtual MAC address as source
- The filtering database on the switch SW-1 will oscillate periodically

- Note that the problem appears despite R1 will not answer to ARP Requests ("Who has 2.2.2.3?") because the incoming path is unavailable
GLBP

- Enhancement (and replacement) of HSRP
- Cisco proprietary
  - Not even available on the entire product line
- Automatic load balancing across default gateways
  - Traffic is distributed across multiple routers
  - No configuration problems (such as in mHSRP) in assigning multiple default gateways to clients and creating multiple groups
- Same first-hop failure recovery capability of HSRP
- A group of routers provides a unique virtual router service
  - One IP address
  - Multiple virtual MAC addresses for forwarding
GBLP: the idea

ARP Request: who has IP 2.2.2.4?
ARP Reply: IP 2.2.2.4 is at MAC 44-44-44-44-44
ARP Request: who has IP 2.2.2.4?
ARP Reply: IP 2.2.2.4 is at MAC 55-55-55-55-55
ARP Request: who has IP 2.2.2.4?
GLBP Functions

- Active Virtual Gateway (AVG)
  - GLBP members elect one router to be the AVG for the group
  - AVG replies to ARP requests for the virtual IP from clients
  - AVG assigns virtual MAC addresses to the active virtual forwarders

- Active Virtual Forwarder (AVF)
  - Each router in the VRRP group (up to 4 per group) routes packets forwarded to its assigned virtual MAC address
GLBP Operation

- An AVG router is elected within each GLBP group
- The AVG allocates a distinct virtual MAC address to each member (the AVFs)
  - In order to move that traffic to another router in case that AVF fails
- If a client ARPs the virtual IP address, the AVG responds with one of the virtual MAC addresses assigned to the AVFs
  - Clients now send their frames to one of the AVFs
- If an AVF fails, another AVF takes over forwarding for that AVF
Load Balancing Algorithm

- Four possibilities
  - None: GLBP operates like HSRP
  - Weighted: each GLBP router in the group will advertise its weighting and assignment; the AVG will act based on that value
    - Used in case the exit links have different capacities
  - Host dependent: this ensures that a host will be guaranteed to use the same virtual MAC address as long as the number of AVFs in the GLBP group is constant
    - Used when using stateful Network Address Translation because it requires each host to be returned the same virtual MAC address each time it sends an ARP request for the virtual IP address
  - Round robin: each AVF MAC address is used sequentially in ARP replies for the virtual IP address
Conclusions

- HSRP/VRRP widely used in practice
  - Simple and effective
- Often, a single group is used
  - Usually outgoing traffic much smaller than incoming traffic
- GLBP
  - Proprietary, not documented
- Take care of L2 issues
  - L2 resiliency
  - L2 flooding
  - Undetected faults