D-Link Switch Training

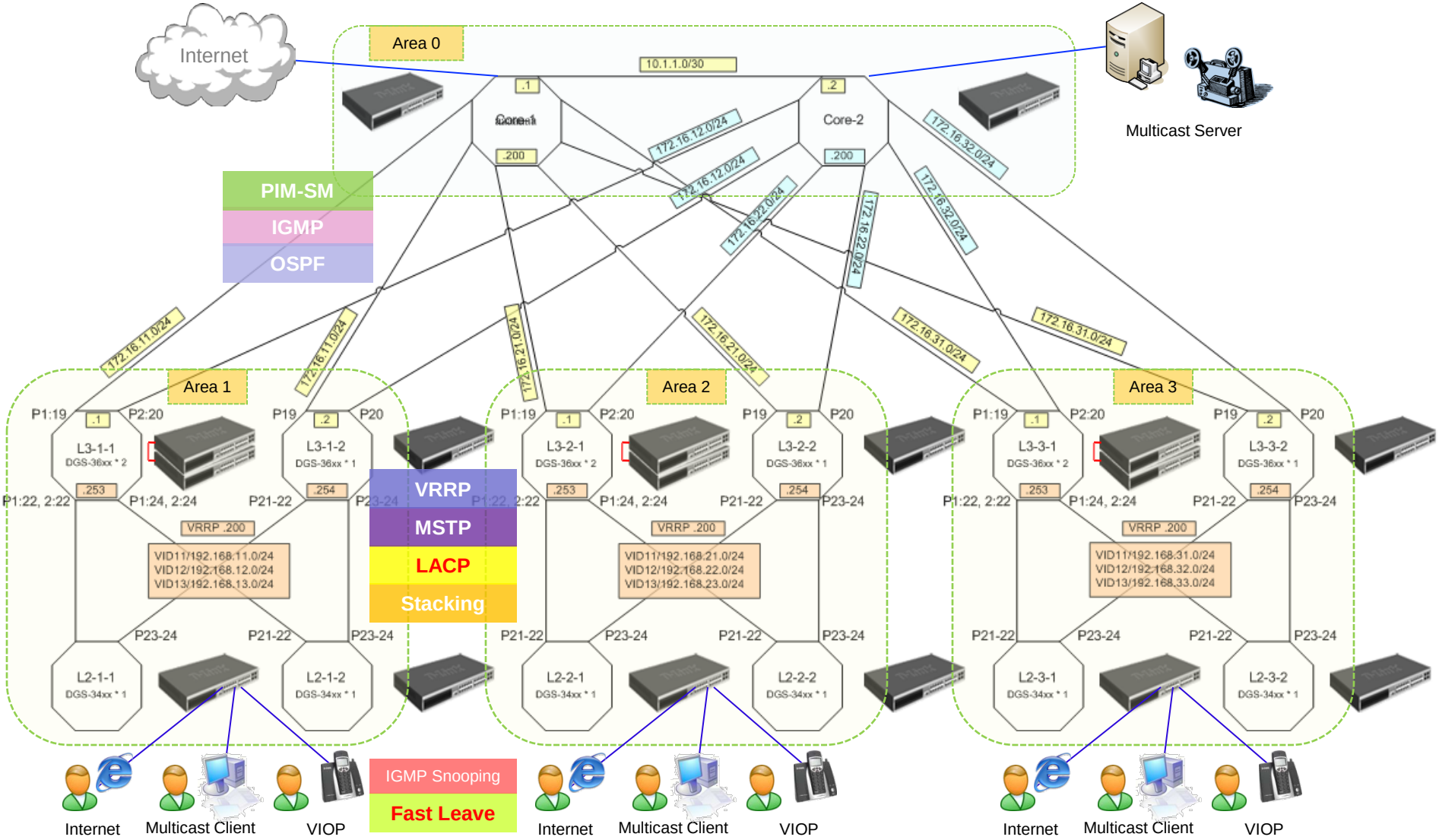
Building Triple Play Network Environment

- Voice, Video, Data
- Key trend in the network market.

Simply requirements of the project :

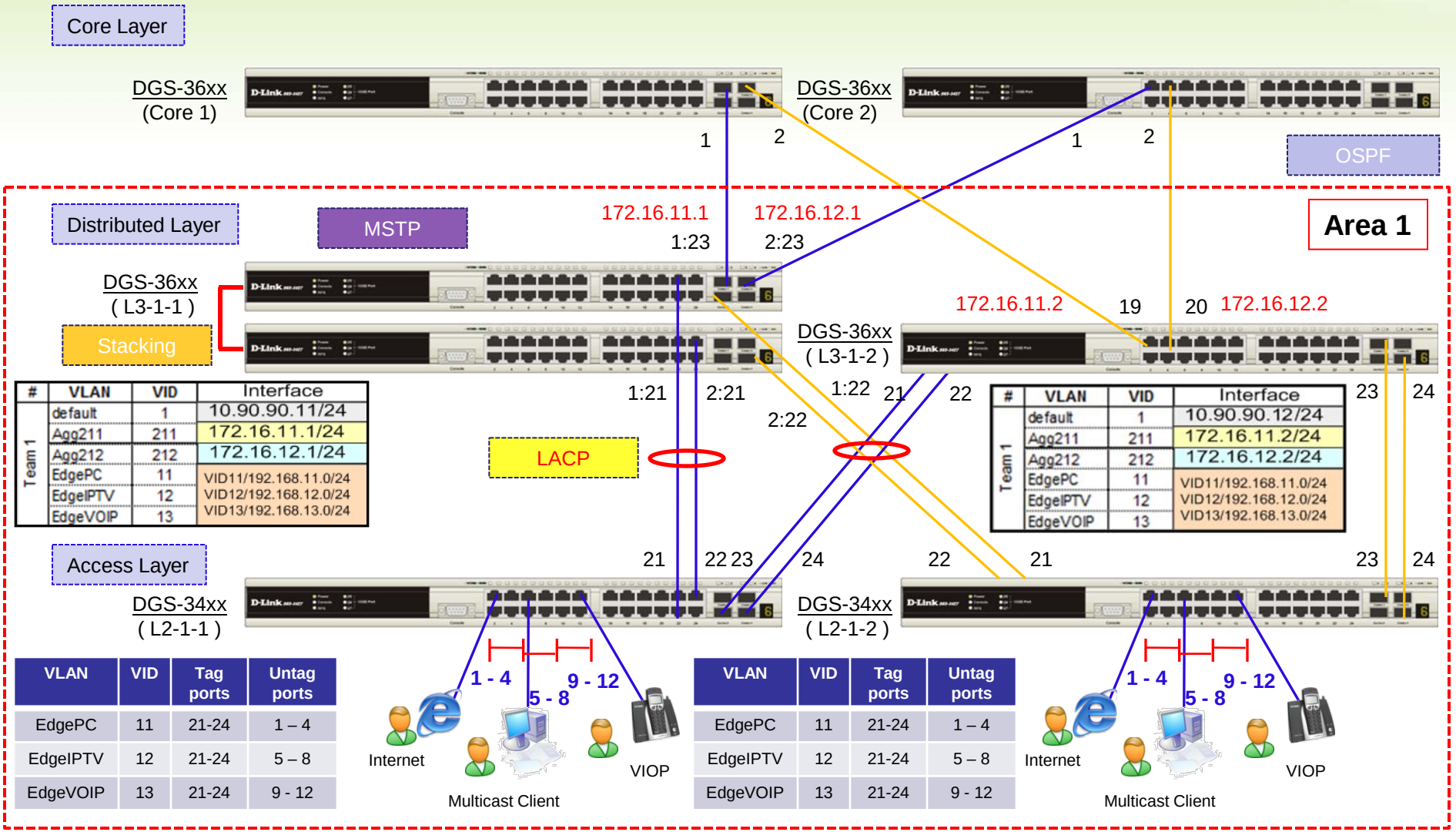
- Voice, Video and Data are all provided in a single access subscription.
- Using Multicast traffic to provide TV channels to end user.
- Need Redundancy Mechanism in the topology design

Triple Play Network Topology



Requirement

Detail network topology



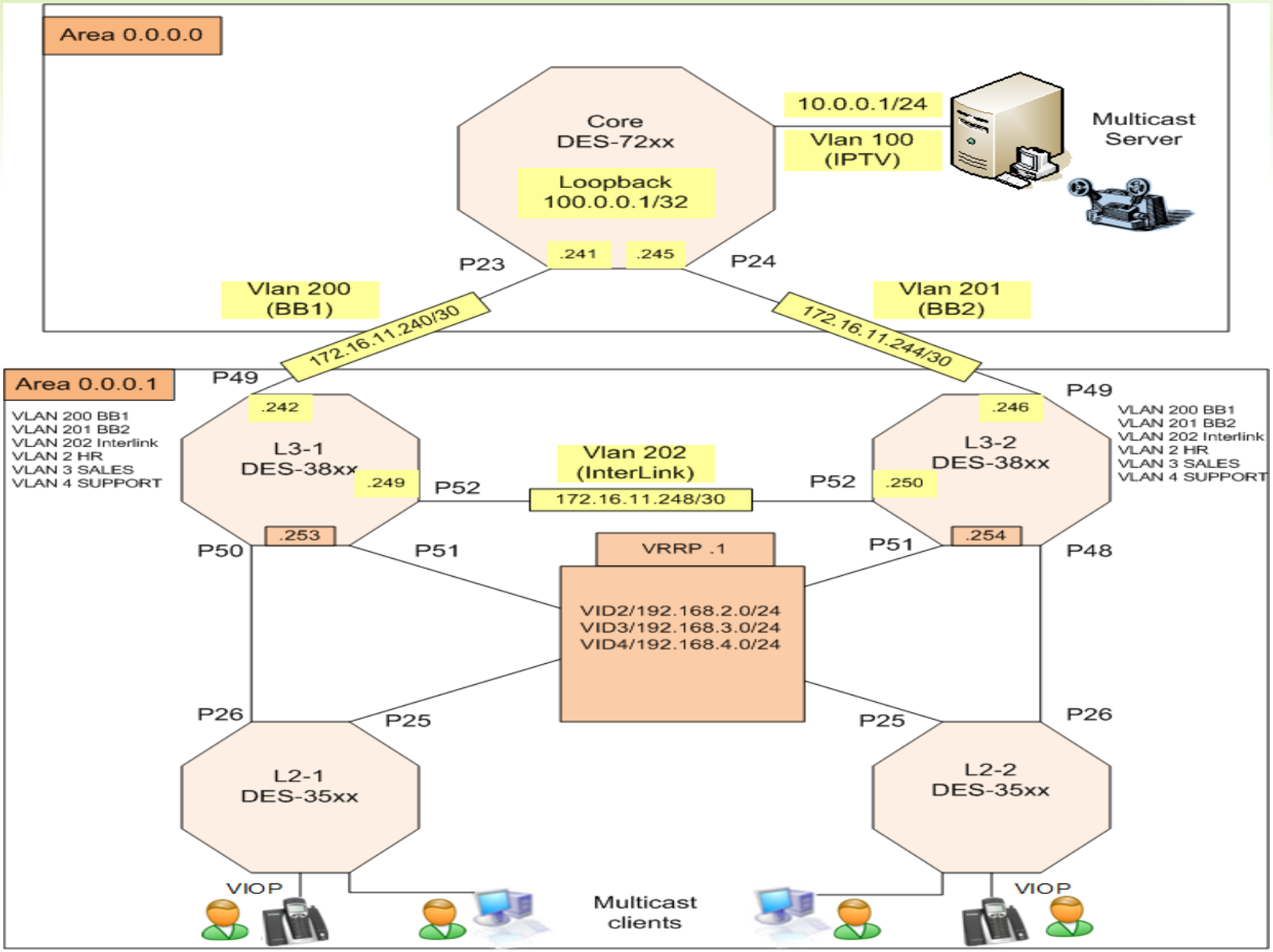
#	VLAN	VID	Interface
Team 1	default	1	10.90.90.11/24
	Agg211	211	172.16.11.1/24
	Agg212	212	172.16.12.1/24
	EdgePC	11	VID11/192.168.11.0/24
	EdgeIPTV	12	VID12/192.168.12.0/24
	EdgeVOIP	13	VID13/192.168.13.0/24

#	VLAN	VID	Interface
Team 1	default	1	10.90.90.12/24
	Agg211	211	172.16.11.2/24
	Agg212	212	172.16.12.2/24
	EdgePC	11	VID11/192.168.11.0/24
	EdgeIPTV	12	VID12/192.168.12.0/24
	EdgeVOIP	13	VID13/192.168.13.0/24

VLAN	VID	Tag ports	Untag ports
EdgePC	11	21-24	1 – 4
EdgeIPTV	12	21-24	5 – 8
EdgeVOIP	13	21-24	9 - 12

VLAN	VID	Tag ports	Untag ports
EdgePC	11	21-24	1 – 4
EdgeIPTV	12	21-24	5 – 8
EdgeVOIP	13	21-24	9 - 12

Network Topology



Prime protocols

- **MSTP** (Multiple Spanning Tree Protocol)
- **OSPF** (Open Shortest Path First)
- **VRRP** (Virtual Router Redundancy Protocol)
- **IGMP** (Internet Group Management Protocol)
- **PIM DM / SM** (Protocol Independent Protocol Dense Mode / Sparse Mode

OSPF
(Open Shortest Path First)

Link State Routing Protocol

- Link = Link between Routers
- State = States of routers

Link State Routing Protocol Characters:

- **Fast convergence** – Affected node respond immediately when network change.

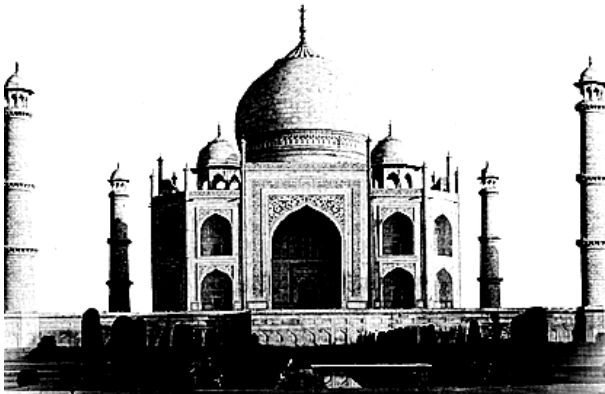
- **Less bandwidth waste** – Sending periodic updates (link-state refresh) at long time internal.

- **Robustness** –

1. Each router maintain its own routing table and independently calculates its best paths to all destinations in the network with Dijkstra's (SPF) algorithm.
2. LSA has sequence number and LSA acknowledge mechanism.

Example: OSPF, IS-IS

Distance-Vector Routing Protocol:
The routers rely on routing decisions from the neighbor Routers do not have the full picture of the network topology. (route by rumor)

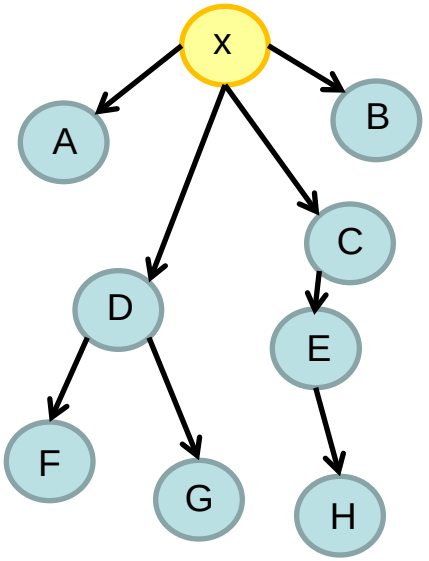


Link-State Routing Protocol Calculation

- Dijkstra's algorithm ➔ calculating the best paths to destinations.

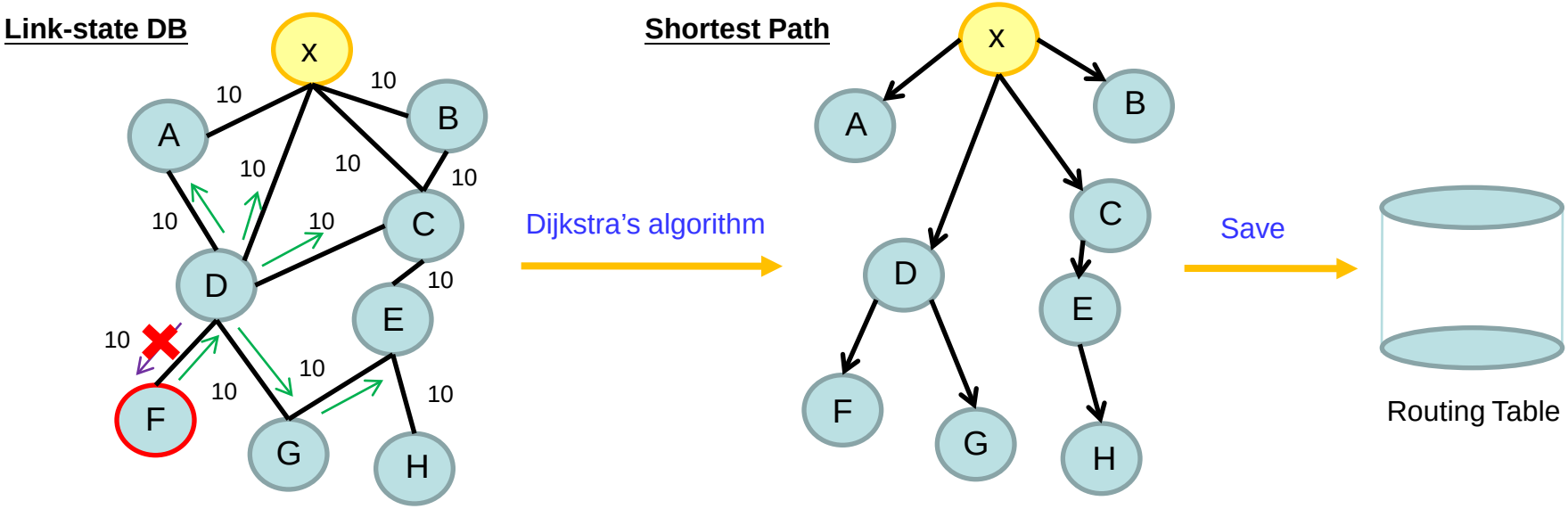
From: Link-State Database

- Every router in an area has the identical link-state DB.
- Each router in the area places itself into the root of the tree that is built.
- The best path is calculated with respect to the lowest total cost of links to a specific destination.
- Best routes are put into the forwarding database (Routing Table) .



Link-State Routing Protocol Calculation (cont.)

- Router F advertises its presence to Router D. Router D passes Router F's and its own advertisements to its neighbors (Router A, X, C and G). Router G passes these and its own advertisements to E, and so on.
- Router D does not advertise Router F's LSAs back to Router F
- Router X has four neighbor routers: A, B, C, and D.
- Each Ethernet link in Figure below is assigned an OSPF cost of 10. By summing the costs to each destination, the router X can deduce the best path to each destination.
- Best routes are put into the forwarding database (Routing Table) .



Assume all links are Ethernet, with an OSPF cost of 10

Link State Routing Protocol

The example of Link-State Routing Protocol operation

1. Detecting network change

2. Create a Link-State Advertisement (LSA) concerning that link and propagates to all neighbor device.

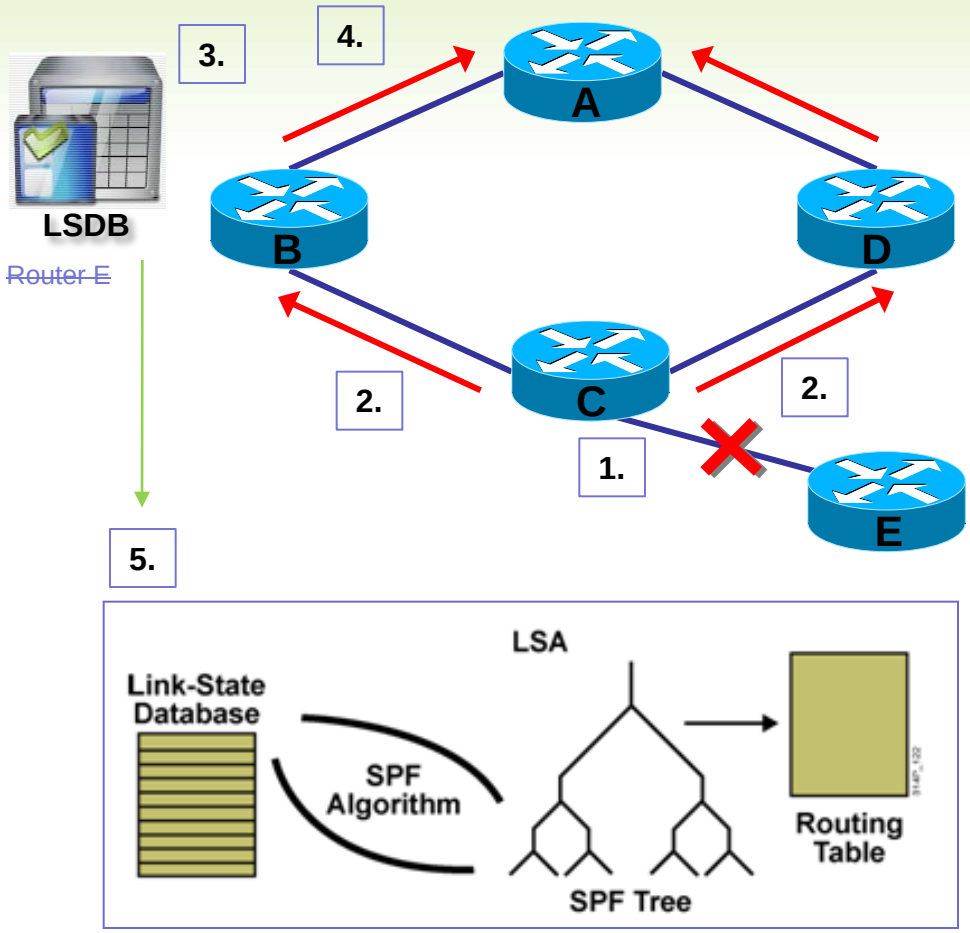
3. Each routing device takes a copy of the LSA, update its link-state database (LSDB)

4. Forwards the LSA to neighboring devices

5. LSDB (Topology Table) is used to calculate the best paths through the network and put it in the routing table.

Drawbacks:

- **Memory Resource Issue**
- **CPU Consumption Issue**



OSPF Overview

OSPF (Open Shortest Path First)

- Link-State Routing Protocol
- Hello / Adjacencies
- Link State Advertisement (LSAs) over all adjacencies ➔ LSDB (Link-State Database)
  - Router's link
  - Router's Interface
  - Router's neighbor
- Flooding LSAs throughout an area / all routers build identical Link-State Database
- SPF(Dijkstra) algorithm to calculate a shortest path ➔ Routing Table

OSPF Router ID

To run the OSPF , a router must have a Router ID.

- A 32-bit unsigned number to uniquely identifies a router in the AS.

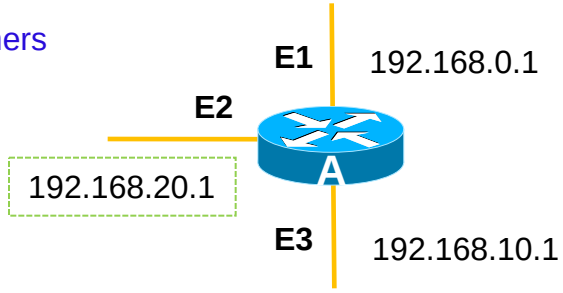
How to generate Router-ID ?

- **Configure manually**
- **The switch / Router select one interface to be Router-ID automatically.**
  - The highest IP address on an loopback interface is chosen by default.
  - The highest IP address on an active interface is chosen.
  - Selection begins at the start of the OSPF process.

Purpose:

- LSDBs use the OSPF router ID to differentiate one router from the others

.



DR/BDR election

DR and BDR election base on:

- **Router Priority –**
- **Router ID**

DR's exist for the purpose of reducing network traffic by providing a source for routing updates,

The DR maintains a complete topology table of the network and sends the updates to the other routers via multicast.

All routers in an area will form a slave/master relationship with the DR. They will form adjacencies with the DR and BDR only. Every time a router sends an update, it sends it to the DR and BDR on the multicast address 224.0.0.6. The DR will then send the update out to all other routers in the area, to the multicast address 224.0.0.5

DR/BDR will be elected by the following rules:

- The router with highest priority value is the DR
- The router with the second highest priority value is BDR
- In case of a tie. The highest Router ID is DR, the second is BDR
- A router with priority of 0 **cannot** be the DR or BDR
- A router that's not DR or BDR is a DROther
- If a router with higher priority comes into the network, it does not preempt the DR or BDR

OSPF_DRBDR election.log

OSPF Adjacency

An OSPF Router transitions a neighbor through several states before the neighbor is considered fully adjacency

Neighbor Discovery

- **Down –**
  - ❖ No Hellos message from the neighbor in the Last RouterDeadInterval.

    If a neighbor transitions to the Down state from some higher state, the link State Retransmission, Database Summary, and Link State Request lists are cleared.
- **Init –**
  - ❖ The Router has seen a Hello message from a neighbor.

Bidirectional Communication

- **Two-Way –**
  - ❖ The Router can see its own Router-ID in the Neighbor field of the neighbor's Hello packet.
  - ❖ DR/BDR election (In the multi-access area)
  - ❖ The Router receives Database Description packet from the neighbor in the init state causes a transition to 2-Way.

Data Synchronization

- **ExStart –**
  - ❖ The Routers establish a master/slave relationship and determine the initial DD sequence number in preparation for the exchange of Database Description packets.
- **ExChange –**
  - ❖ The router sends Database Description packets describing its entire Link-State database to neighbors that are in the Exchange state. The router may also send Link State Request packet for requesting more recent LSAs.

Full Adjacency

- **Loading –**
  - ❖ The router sends Link State Request packet to neighbors.
- **Full –**
  - ❖ Neighbors in this state are fully adjacent..

OSPF Areas

OSPF Areas

Problem: **In the Link-State Routing Protocol, all router must keeps all routing information in the LSDB**

- The large scale network cause the need of larger LSDB ➔ Memory Issue
- Dijkstra (SPF) calculation comparing all of these possible routes can be very complex and take significant time ➔ CPU Issue

Solution: Area ➔ Reduce the impact on the CPU / Memory.

Link-State routing protocols use a two layer areas.

1. Transit area –

1. Fast and efficient forwarding IP packets.
2. Transit area interconnect with other OSPF area types.
3. OSPF area 0 / backbone Area
4. summaries the topologies of each area to every other area.

2. Regular area –

1. Connect users and resources
2. All traffic from the regular area must cross a transit area.
3. OSPF area **not** 0 / many are types

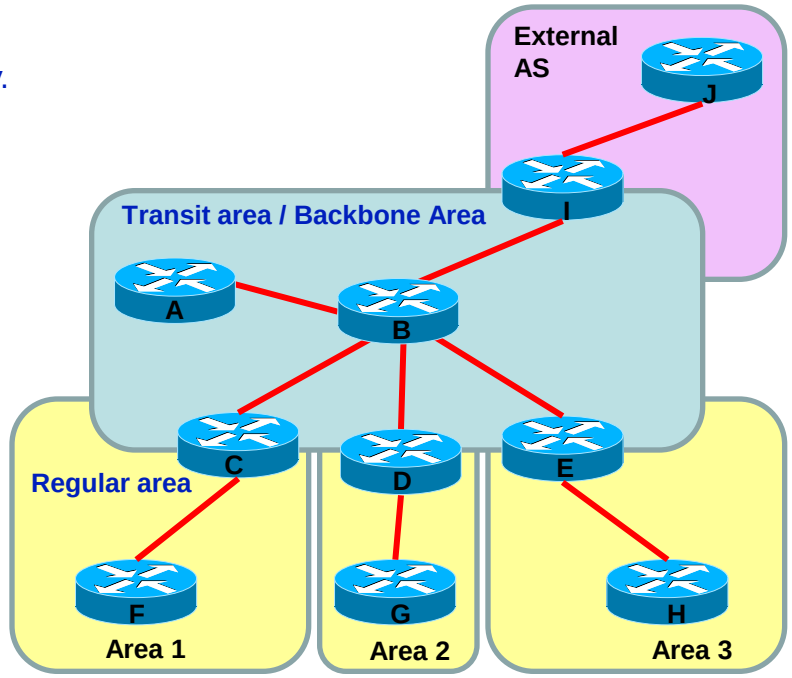
The OSPF area characteristics:

- **Smaller link-state database**
- **Routers in the area share an identical link-sate database.**
- **Reduced link-state update(LSU) overhead**

Detailed LSA flooding stops at the area boundary

- **Requires a hierarchical network design**

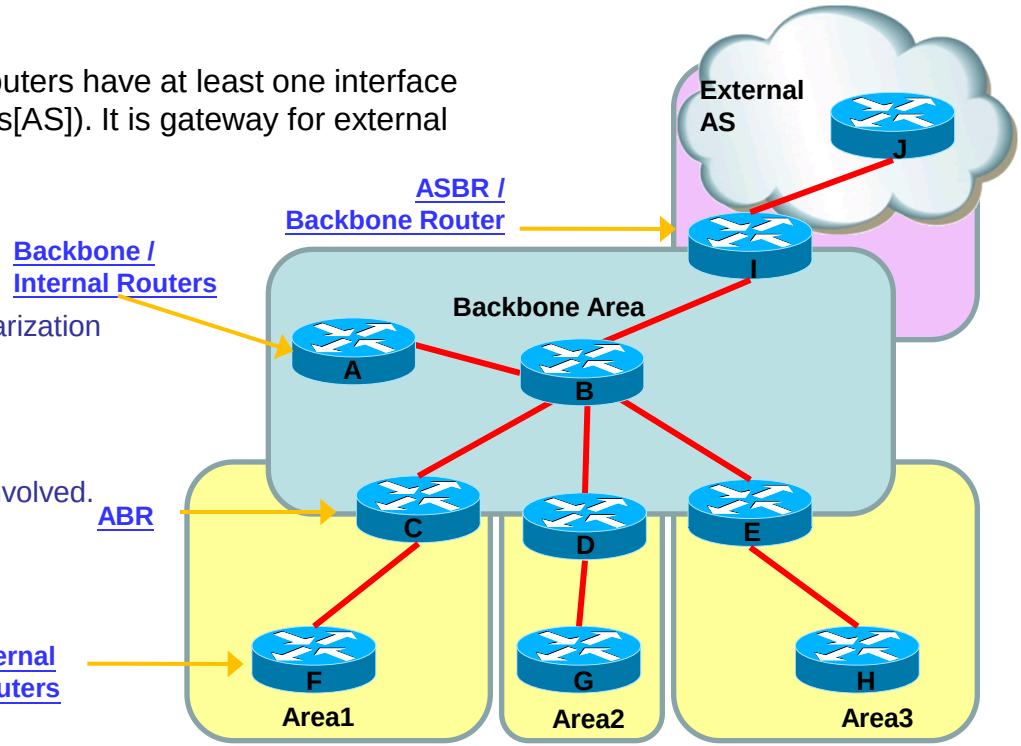
50-100 routers per area. (Cisco advice)



Example of Area ID
1. 0 = 0.0.0.0 (reserved for Backbone)
2. 275 = 0.0.1.19
275 ➔ 100010011 ➔ 00000001 00010011 ➔ 1. 19

Type of OSPF Routers

- **Internal Router –** Router's interfaces all are in the same area.
- **Backbone Router –** Routers have at least one interface connected to area 0
- **ABR (Area Boarder Router) –** Routers' interface connect to one or more areas to the backbone and act as a gateway for inter-area traffic.
- **ASBR (Autonomous System Boundary Router)–** Routers have at least one interface attached to an external internetwork (another autonomous[AS]). It is gateway for external traffic, injecting routs into the OSPF area.
  - It separates LSA flooding zone.
  - It becomes the primary point for area address summarization
  - It functions regularly as the source of default routers
  - It maintains the LSDB for each area with which it is involved.
- A router can exist as more than one router type.
- A router has a separate LSDB for each area to which it connect



VRRP

(Virtual Router Redundant Protocol)

VRRP – (Virtual Router Redundant Protocol)

Issue: Achieve higher network reliability.

Normal configuration:

- **Step 1:** manually configure the IP-Address / Mask / Default Gateway.
- **Step 2:** set default gateway to point to Router A.
- **Step 3:** Router A will find the destination and forward packets.
- ***Problem:*** The client would be isolated from the external network.

i.e. If the router fails, the connection is broken.

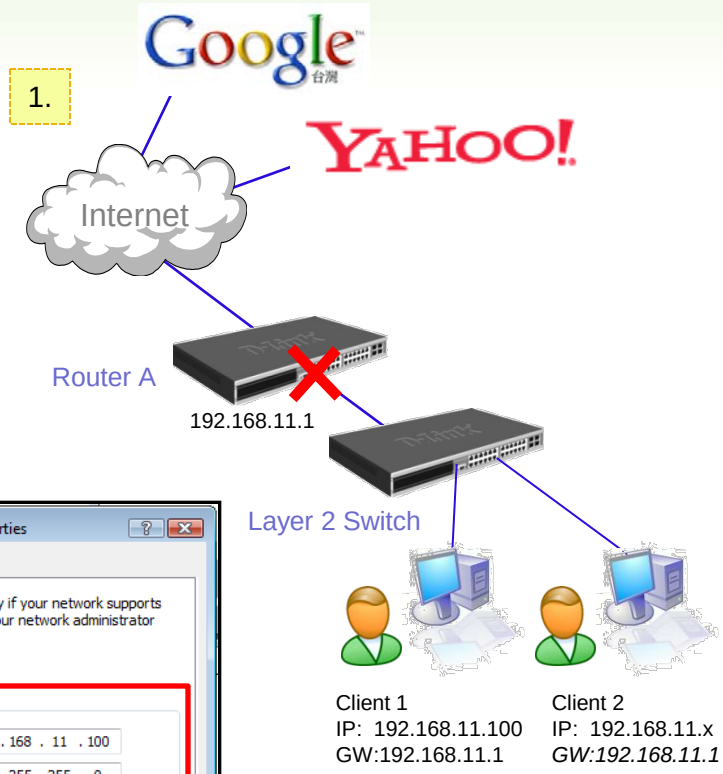
What to do?

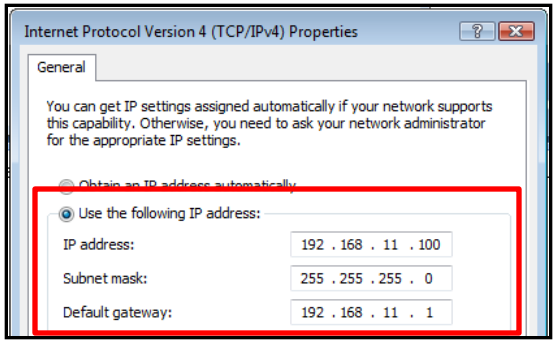
a. Add another router in the network ?

b. Run dynamic routing protocol (RIP, OSPF) ?

The solution:

VRRP makes it easy to settle this problem.



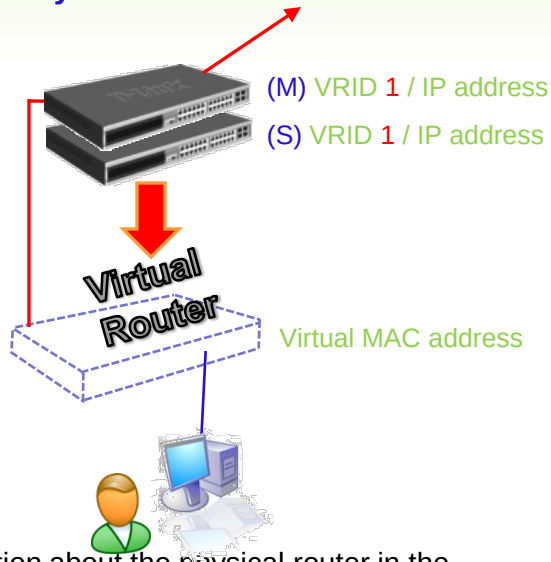


Network topology without VRRP

VRRP – (Virtual Router Redundant Protocol)

- **VRRP is a fail-over protocol designed for LAN.**
- **Provides communication continuity and reliability if the original default gateway is broken.**

How to run VRRP:



■ The Virtual Router Identifier (VRID) and IP address should to configured on each router.

■ When a client communicates with the virtual router it does not need to have any information about the physical router in the network.

■ After the VRRP election process, one of the routers is Master, which is then responsible for transmitting traffic. The other router(s) will be backup. The Virtual IP address is associated with the virtual router.

- **Priority**
- **IP Address (Higher IP Address wins)**

VRRP – (Virtual Router Redundant Protocol)

■ VRPP defines one type of packet only. (VRRP packet). It is multicast packet (224.0.0.18).

- The master router sends VRRP packet periodically to check parameters of the virtual router.
- To select master router.

Example of VRRP:

■**Step 1:** VRRP groups Router A and Router B into a Virtual Router.

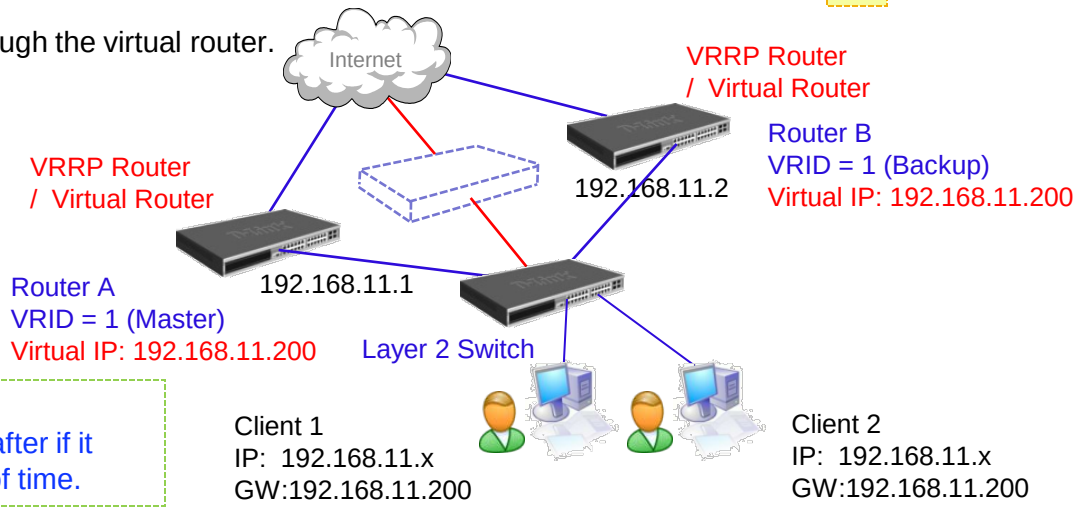
■ **Step 2:** The virtual router has its own IP address of 192.168.11.1, which may the same as interface address of certain router, called IP address owner.

■ **Step 3:** Attached IP Address (192.168.11.200) to Virtual Router.
The clients don't care the physical interface of Router A and B. (11.1 and 11.2)

■ **Step 4:** Client 1 and 2 set 192.168.11.200 to be Default Route.

■ **Result:** The clients communicates with internet through the virtual router.

2.

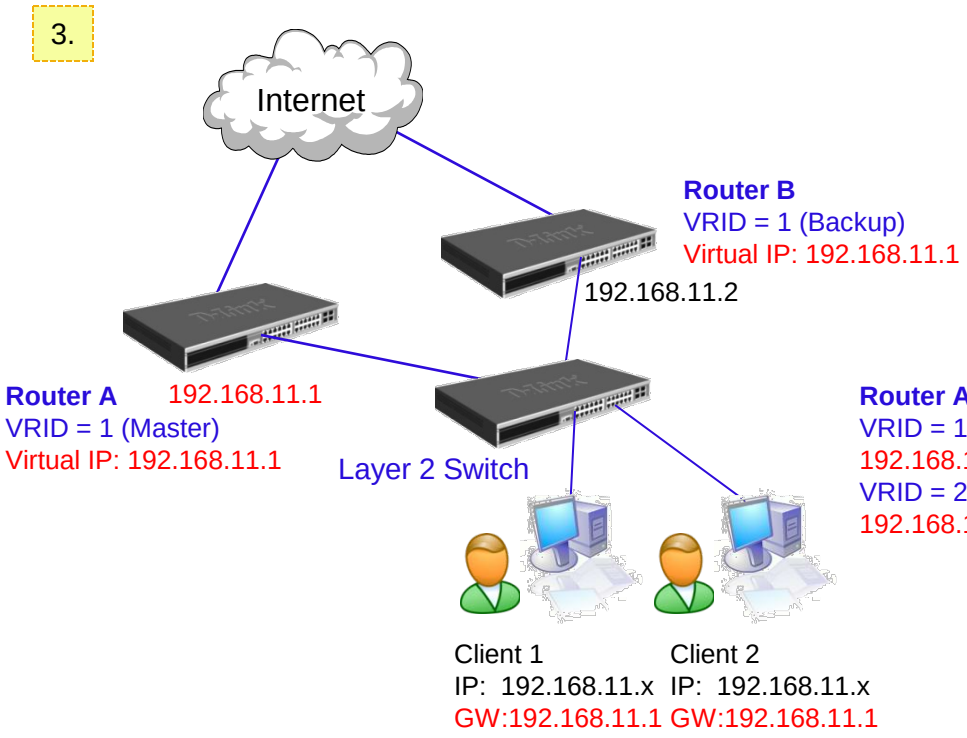


Backup Router → Master :
The Backup Router will transfer the state into Master after if it has not received packet from the master for a period of time.

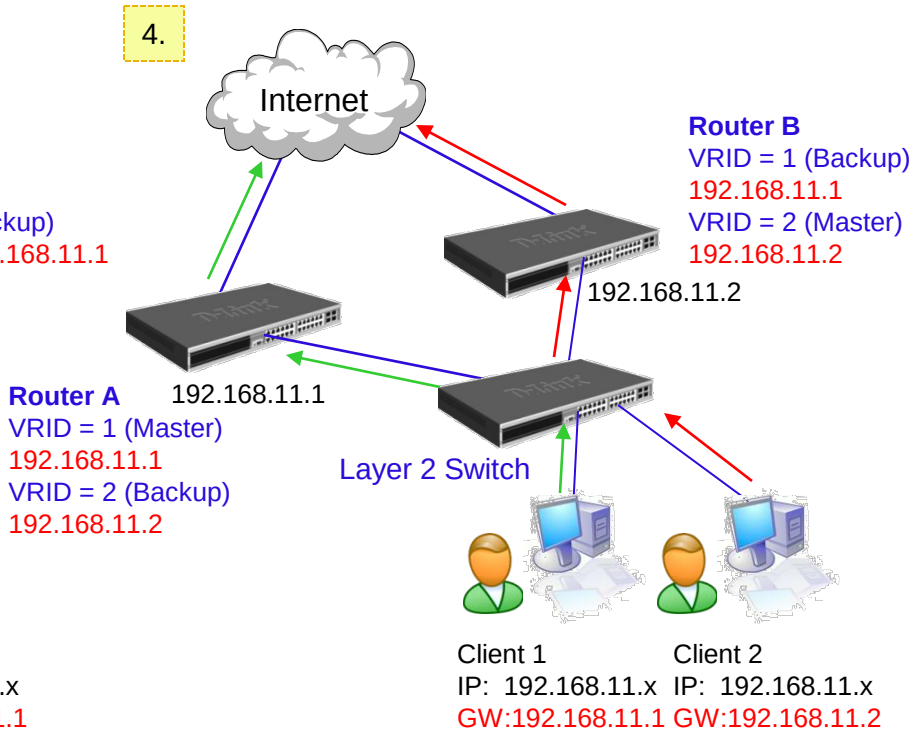
VRRP Example – (Virtual Router Redundancy Protocol)

3. Use the real interface to be Virtual IP.

4. One router is allowed to backup for multiple virtual routers. Router A is the master of virtual router 1 and it's backup router of virtual router 2 at the same time. Client 1 points to router A as the default gateway and Client 2 points to Router B as the default gateway. This is known as **Load sharing**.



Network topology with VRRP

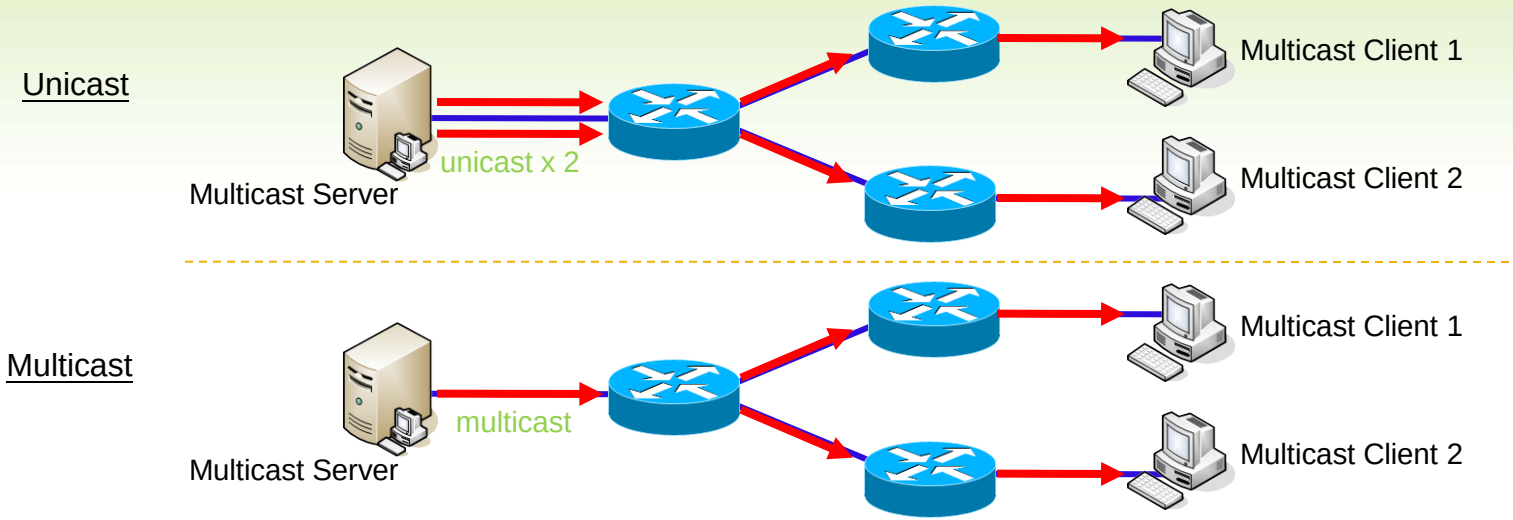


Network topology with VRRP providing load splitting

IGMP

(Internet Group Management Protocol)

Multicast Traffic



Multicast traffic can choose following method:

- **Unicast (Multiple copies, Multiple receivers) –**
  - an application sends two copies of traffic to two clients.
  - Wastes bandwidth

- **Multicast (Single copy, Multiple receivers) –** the most efficient solution is multicast.
  - The client device decides whether or not to listen to the multicast address
  - Forcing the network to forward packets only when necessary.

- **Broadcast (Single copy, All receivers) –** an application sends only one copy of each packet using a broadcast address.
  - Eeach host device must process the broadcast data frame.
  - Cannot pass through the router.

Multicast Advantage/Disadvantage

Advantage :

- **Enhanced Efficient** ➔ Single multicast transmission utilizes network bandwidth more efficiently.
- **Optimized Performance** ➔ Less data requiring forwarding and processing.

Disadvantage :

- Uses **UDP** as the transport layer protocol. ➔ (not reliable, no ACK mechanism)
- No Congestion Avoidance ➔ (ex: TCP windowing and pause frame)
- Duplication: some multicast protocol mechanisms (ex: Asserts, Registers and Shortest-Path Tree Transitions) cause the occasional generation of duplicate packets.

IGMP – (Internet Group Management Protocol)

- The current version of IGMP:
  - IGMP version 1 (RFC1112)
  - IGMP version 2 (RFC2236)
  - IGMP version 3 (RFC3376)

- The IGMP manages multicast group memberships mainly based on:
  - How a client **Join (Report)** a group.
  - How a client **Leave** a group.
  - How a router **Query** clients.

- Hosts ➔ use IGMP to dynamically register themselves in a multicast group on a particular subnet.

- Router and multilayer switches ➔ keep listening to IGMP message and periodically send out queries to discover which groups are active or inactive on particular subnet or VLAN.

Multicast Group ex: 225.1.1.10

Channel List.

CH 10	HBO Movie Channel	**Group 1:** 225.1.1.10 Movie Channel
CH 11	CNN.com News Channel	**Group 2:** 225.1.1.11 News Channel

IGMP v1

:: Query Mechanism ::

- The Querier send IGMP Query to all clients (224.0.0.1) periodically (60 sec.) and TTL value of packet is equal to 1.

- **No Querier election mechanism.** The designated router (DR) elected by a multicast routing protocol (ex: PIM).

:: Join Mechanism ::

The timing of sending Join/Report Message :

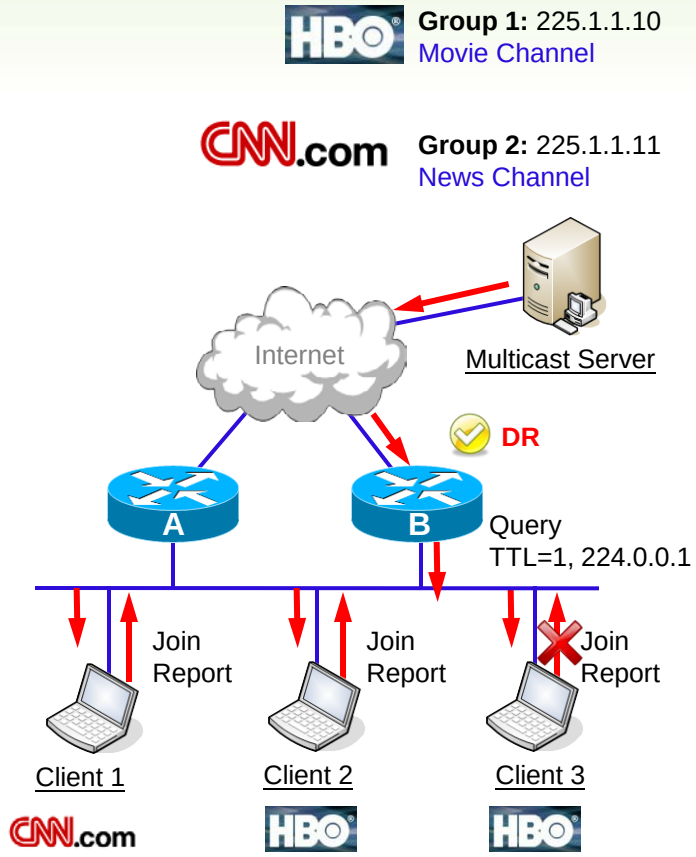
1. When receiving an IGMP query message, clients responds with IGMP **Join Report** for the group it interest in.

2. When a host want to join a multicast group, it send out a multicast membership report to the router.

- **Report Suppression Mechanism** – If a client receives a given group report (ex: 225.1.1.10) from other member, it will keep quite and will not send the same report to ask multicast traffic.
  Benefit : Reduce bandwidth over the local subnet.

:: Leave Mechanism ::

- Clients leave multicast group quietly without sending notification to the multicast router. The multicast router stop forwarding traffic after client response timeout (no client in a group).



IGMP Version 2

IGMP v2

- **IGMP v2 solves the limitation (no leave mechanism) of IGMP v1.**
- **Backward compatible with IGMP v1**

Add two features:

- **Querier Election Mechanism**
- **Leave Group Message**

Host sends leave message if it leaves the group and is the last member.
(reduces leave latency in comparison to v1)

IGMP v2

:: Query Mechanism ::

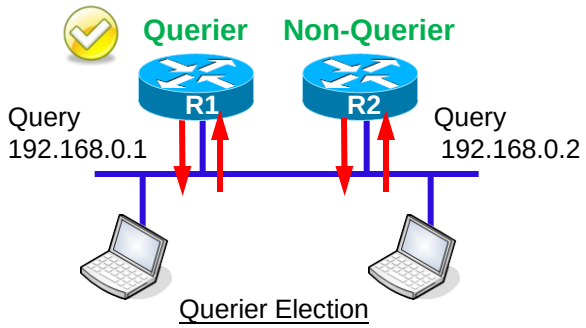
- Query send with multicast IP address (224.0.0.1) and have an IP TTL equal to 1.
- Query interval is 60-120 sec (default is 60 sec.).
- **Query Election mechanism** ➔ Resolves multiple queries on single multicast subnet. (IGMP v1 has no this mechanism).
- **Group Specific Query.** ➔ be aim at a specific group to query.

❖ **Querier Election Mechanism**

- **Step 1 :** Initially, IGMP v2 routers come up and think themselves as queriers and send a IGMP general query message.
- **Step 2 :** When a IGMP router receives a query message with lower source IP than itself, it becomes non-querier.
- **Step 3 :** The IGMP routers with **lowest** IP address will be elected as the Querier.
- **Step 4 :** After election process, all non-querier routers start a timer, known as "other querier present timer". If a router receives a query before the timer expire, it will reset the timer. Otherwise, it assumes the querier fails and re-initiates a election process.

:: Join Mechanism ::

- A client can send the join packet **any time** and doesn't wait for receiving a query message in order to reduces join latency. (Same as IGMP v1, Asynchronous Join)
- Suppresses mechanism. (Only one member per group responds with a report to a query.)



IGMP v2

:: Leave Mechanism ::

- **Leave group mechanism**

Step 1: A Client sends the Leave message to all routers (224.0.0.2) on local subnet.

Step 2: When receiving the “Leave message”, the querier feedbacks number of group-specific queries to the associated group. In order to confirm if there are still other clients wishing to receive traffic for the group.

Step 3: One of the remaining members of the group should response a join report within the maximum response time (Query-Interval Response Time) set in the query message.

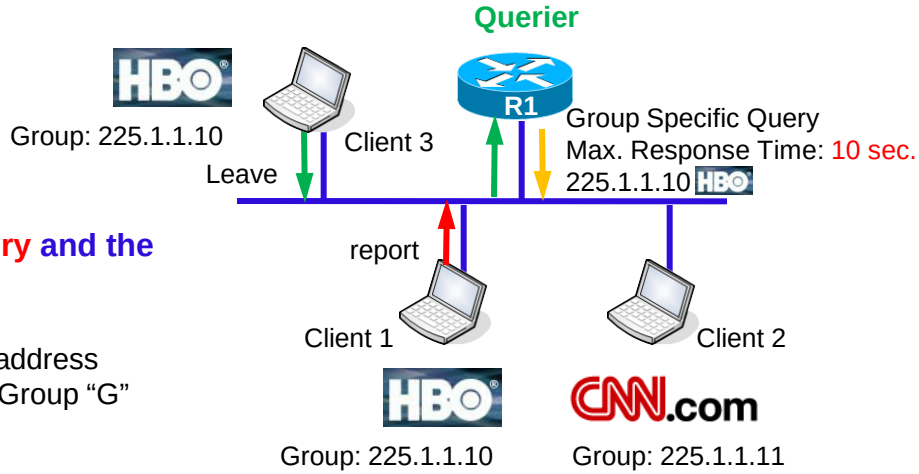
Step 4: If the querier receives join message sent by a client, it will keep send traffic into the subnet. Otherwise, the querier will assume no client interest in the group and stop forwarding the traffic into the subnet

- **The benefit of Group Specific Queries**

a. Quickly find out if any members are left in a group.
b. Router doesn’t need to ask all groups for a report.
c. Shorten the time of stopping flood the traffic.

- **The difference between the Group Specific Query and the General Query:**

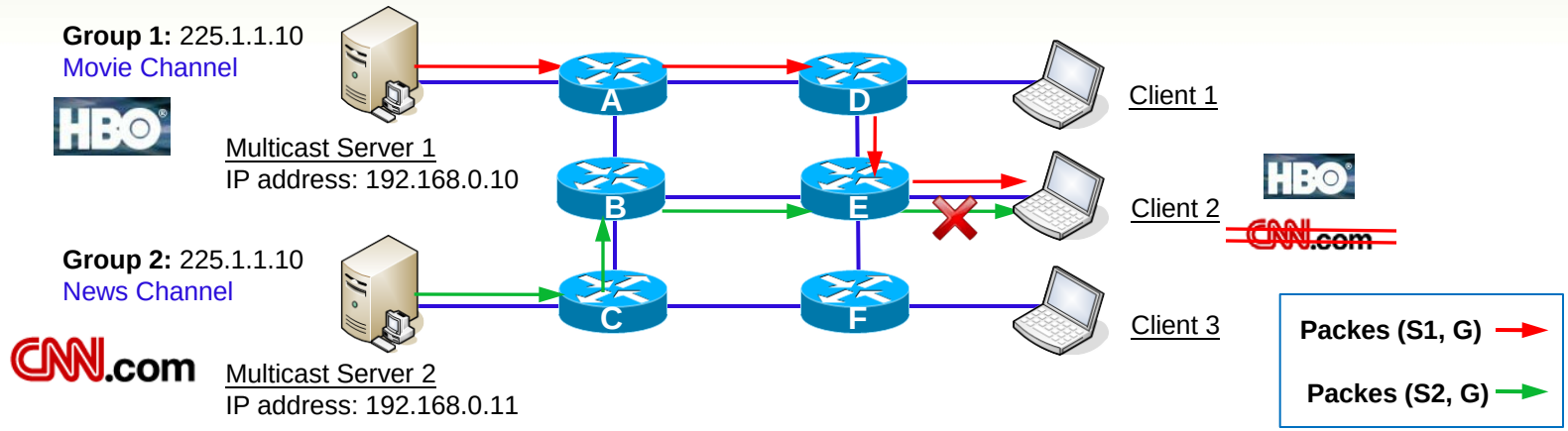
General Query ➔ multicast to the “All-Hosts” (224.0.0.1) address
Group Specific query for Group “G” ➔ multicast to the Group “G” multicast address.



IGMP Version 3

IGMP v3

- **RFC 3376**
- **Enhance host control capability using Source Filter Mode (Include/Exclude Source Lists)**
  - To allow hosts to receive/reject a designated multicast group from **one or a set** of multicast servers.



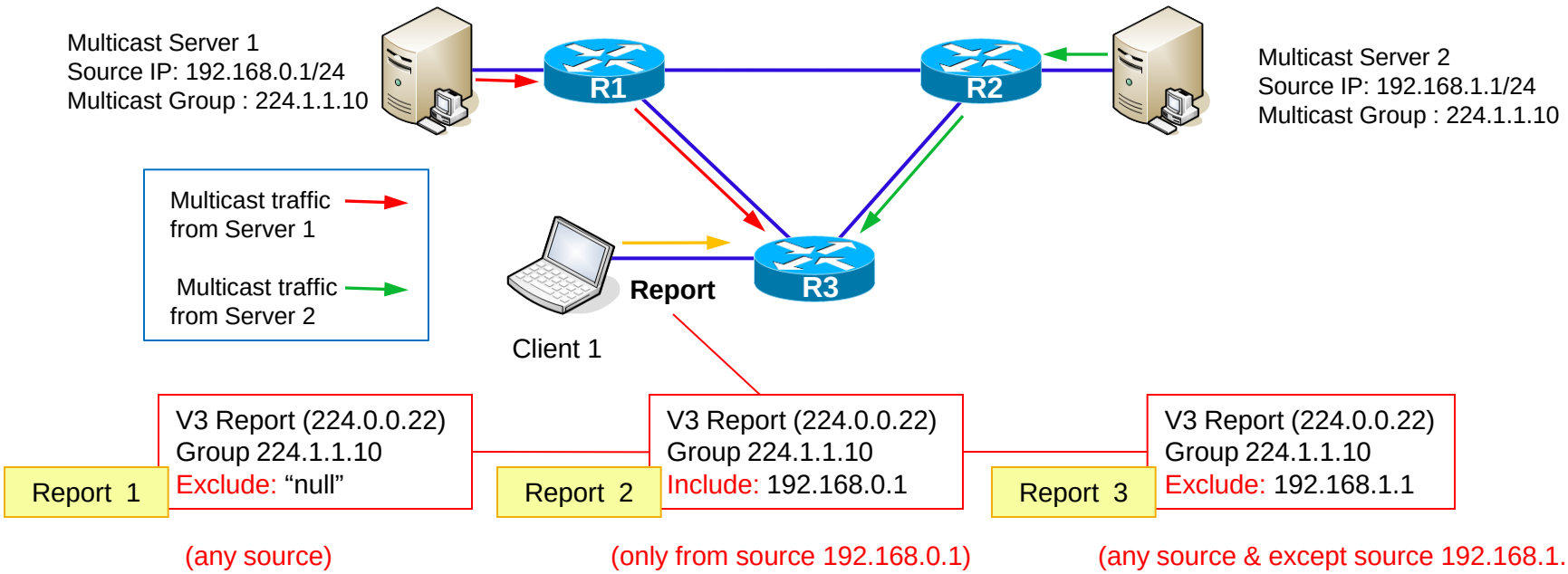
  - Ex: If Client 2 only want to see movie channel (HBO), it can just include the server 1 into its report.

- **Enhance query and report capabilities.**
  - **Group and source-specific-queries**
  - **General query** ➔ multicast to the "All-Hosts" (224.0.0.1) address and **doesn't** carry group address and source address.
  - **Group specific query** ➔ multicast to the Group "G" multicast address and **carry** a group address, **no** source address.
  - **Group and source specific query** ➔ multicast to the Group "G" multicast address and **carry** a group address and one or more source address.

IGMP v3 Join Example

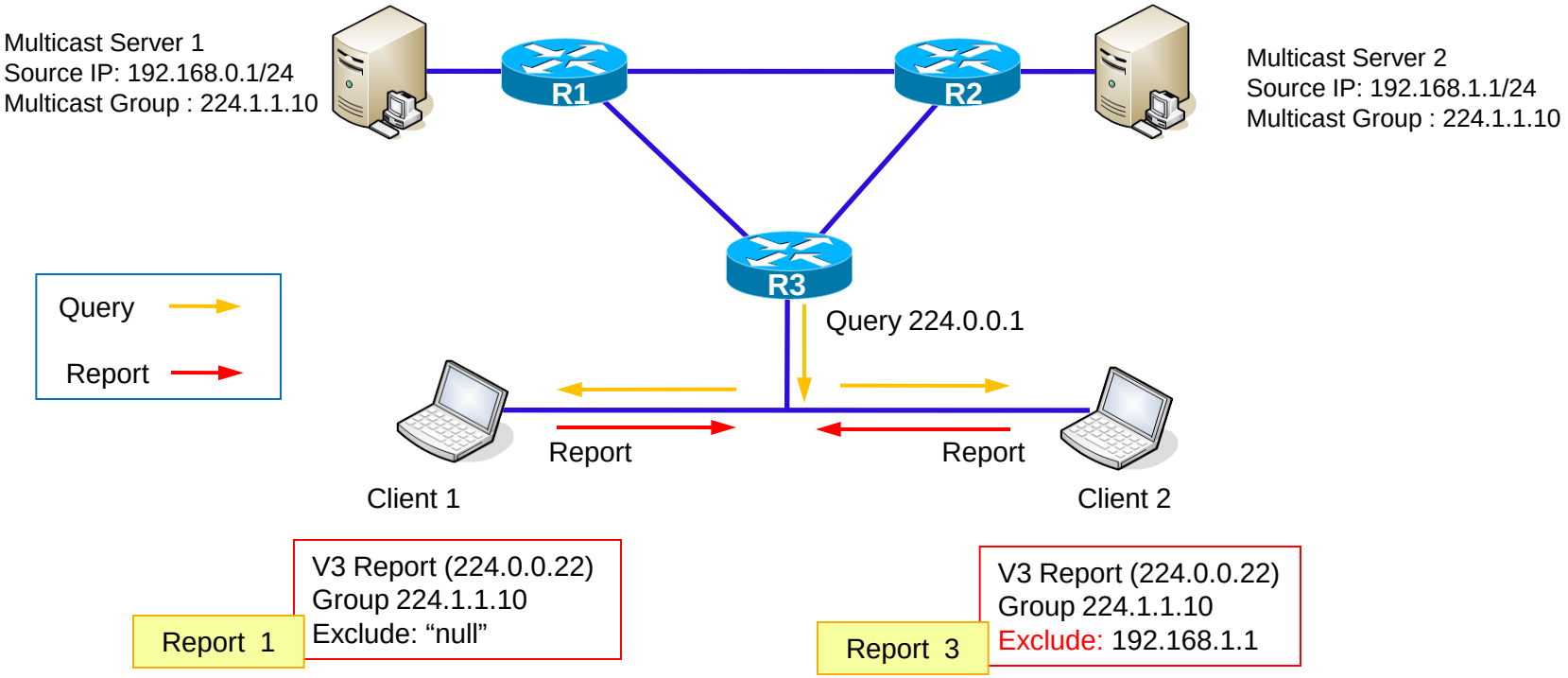
IGMP v3 Join Example

- **Report 1** ➔ Client 1 sends a report to join all source of the multicast group 224.1.1.10.
- **Report 2 (Joining only specific Sources / Include)** ➔ Client 1 sends a report to join only the source (192.168.0.1) of multicast group 224.1.1.10.
- **Report 3 (Joining only specific Sources / Exclude)** ➔ Client 1 sends a report to join all source of the multicast group 224.1.1.10 except the group from the source (192.168.1.1).



IGMP v3 Maintaining State

- **IGMP v3 Maintaining State**
  - No Report Suppression mechanism.
  - The router multicast periodic membership Queries to the “All-Hosts” (224.0.0.1) group address.
  - All hosts responded by sending back an IGMP v3 Membership report that contains their specified multicast address list for the interface.



Category	Function	IGMP v1	IGMP v2	IGMP v3
Query	Periodically Query	Yes 224.0.0.1 TTL = 1 Interval = 60-120 (60)	Yes	YES
	Group-Specific Query	No	Yes	Yes
	Group-and –Source Specific Query	No	No	YES
	Query Election mechanism	No	Yes	YES
Report	Report suppression	Yes	Yes	No
	Asynchronous report	Yes	Yes	YES
Leave	Leave Notification	No leave quietly	Yes	YES
	Include/Exclude Mechanism	No	No	YES
IGMP v2 use IGMPv1 membership report for backward-compatibility with IGMP v1				

PIM DM/SM

(Protocol Independent Multicast Dense Mode / Sparse Mode)

Outline

- PIM Overview
  - PIM-DM (Dense Mode)
  - PIM-SM (Sparse Mode)

PIM Overview

- **PIM (Protocol Independent Multicast) –**
  - Provides IP multicast forwarding **based on any unicast routing protocol** (ex: OSPF, RIP).
  - When a multicast packet arrives on an interface of router, it should execute **RPF (Reverse Path Forwarding mechanism)** to implement multicast forwarding.

- **PIM has two modes based on forwarding mechanism :**
- **:: Dense-Mode ::**
  - Uses **"Push"** Model – Assumes that **at least one multicast group client** on each subnet of the network.

  Step 1: Routers **flood** multicast traffic throughout all the network.

  Step 2: Routers **prune** back where it has no client interesting in multicast traffic.
  - Flood & Prune behavior (typically every 3 minutes)

- **:: Sparse-Mode ::**
  - Uses **"Pull"** Model – Assumes that **no receivers** interest in multicast traffic unless a client ask for it.
  - Uses a Rendezvous Point (RP) – sender and receiver "rendezvous" at this point to learn each other.
    - Sender are "registered" with RP by first-hop router.
    - Receivers are "Joined" to the Shared Tree (root is RP) by their local Designated Router.

Multicast Forwarding – Reverse Path Forwarding (RPF)

- **Unicast Routing** ➔ where the packet is going. **(Destination)**
- **Multicast Routing** ➔ where does the packet came from. **(Source)**

How to check source?

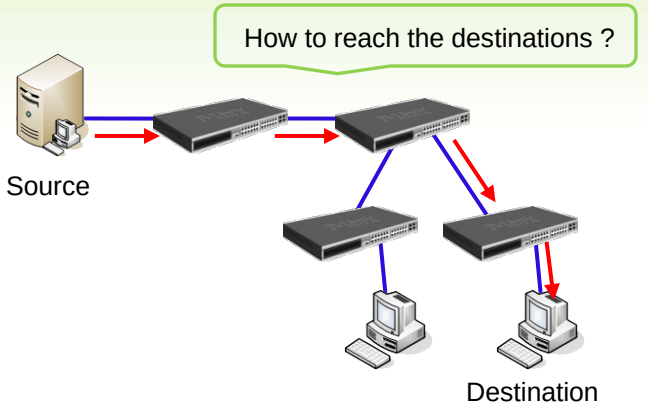
Multicast Routing uses Reverse Path Forwarding (RPF check).

What is RPF?

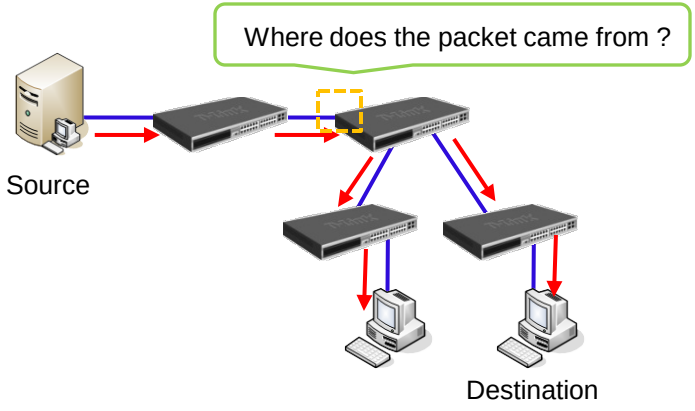
1. The check mechanism to determine that router should **forward** or **drop** packets according to the packets coming from which interface.

2. RPF is a key point in multicast forwarding.

3. Prevent duplicate forwarding issue.

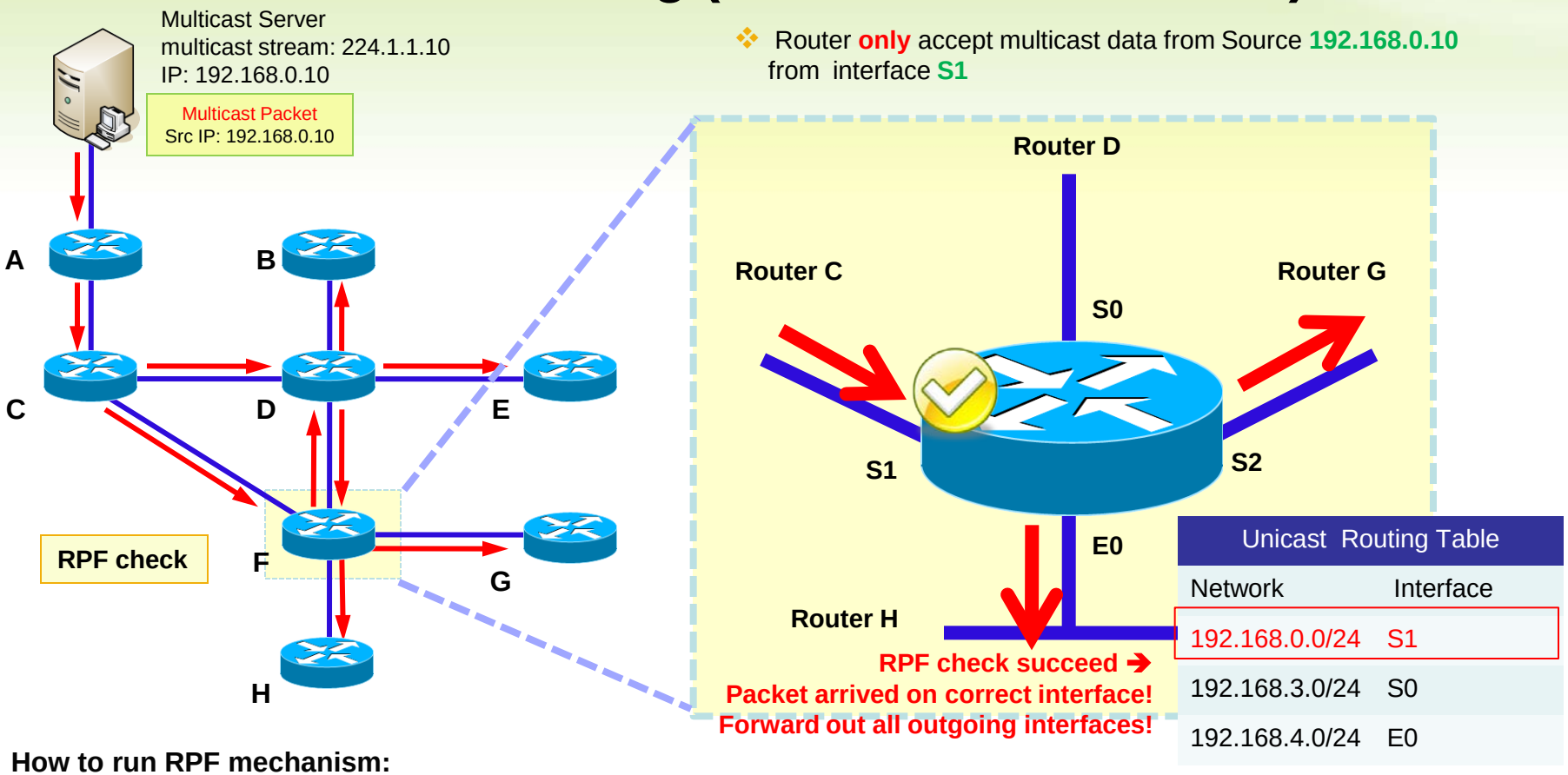


Uni-cast Routing



Multicast Routing

Reverse Path Forwarding (RPF Check Succeeds)



❖ Router **only** accept multicast data from Source **192.168.0.10** from interface **S1**

Unicast Routing Table	
Network	Interface
192.168.0.0/24	S1
192.168.3.0/24	S0
192.168.4.0/24	E0

How to run RPF mechanism:

❖ **Step 1:** Take out source IP address of multicast packets and check unicast Routing table to determine whether the packets arrived on correct interface.

❖ **Step 2:** If the packet has arrived on the interface leading back to the source, the RPF check is successful, and then the router replicates and forwards the packet to the outgoing interfaces.

❖ **Step 3:** If the RPF check fails, the router drops the packet silently.

Rendezvous Points –

Static RP
Bootstrap Router

PIMv2 Static RP / BSR Overview

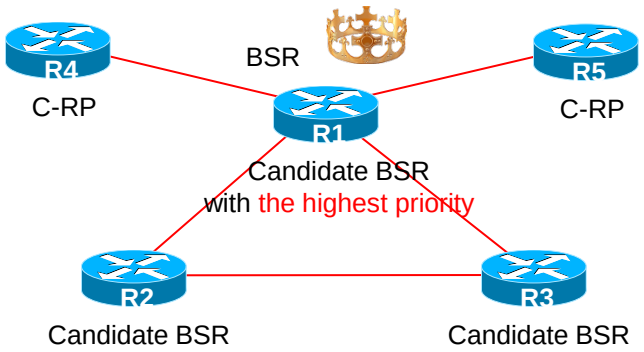
- **RP is a important concept in PIM-SM.**
- **In small-size, simple network topology** – One RP is enough to cover all multicast information / traffic handling.
- **In large scale network environment** – Need more RPs to shard the loading and optimize the topological structure of the RPT.

:: Static RP ::

- **Suitable for small-size network topology.**
- **It must configured on every router and all routers should point to the same RP address.**
- **No RP Fail-over**

:: BSR (Bootstrap Router mechanism) ::

- **Suitable for large scale network environment network topology.**



The roles in BSR mechanism:

- **Candidate BSR (C-BSR) –**
  - A network can contain one or more routers serves as Candidate BSR (C-BSR).
  - BSR will be elected from those Candidate BSR.

- **Bootstrap Router (BSR) –**
  - The BSR is elected from a collection of Candidate BSR's.
  - If the current BSR fail, the new election is triggered to avoid service interruption.
  - Bootstrap router collects all C-RP announcements into a database (RP-set) and periodically send the RP-set out to all other routers in the network.

D-Link
Building Networks for People

greenethernet

Q & A