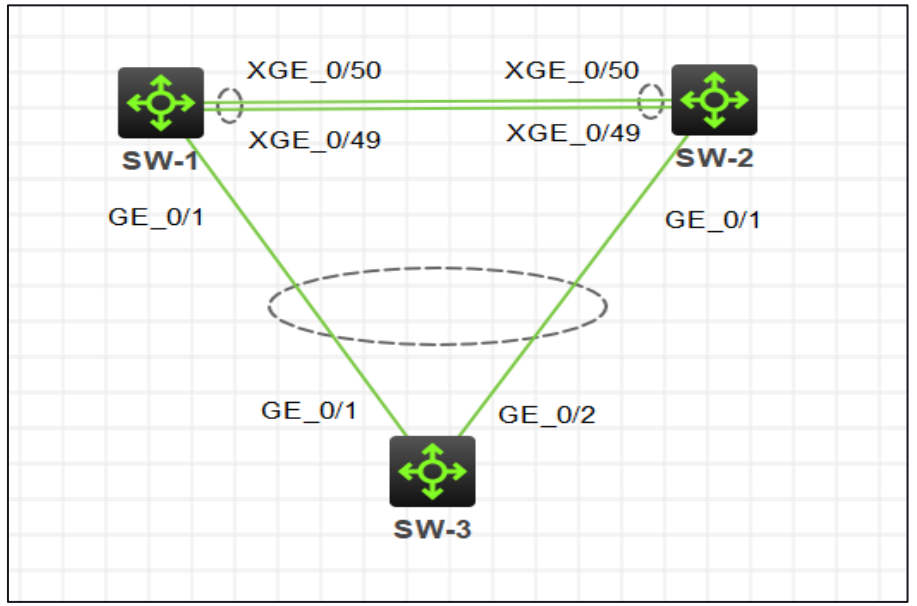
Layer 2 Technologies Advanced

Lecture 1



SoftUni Team

Technical Trainers

Software University

SoftUni

Software University

https://softuni.bg

Table of Contents

Have a Question?

Software University

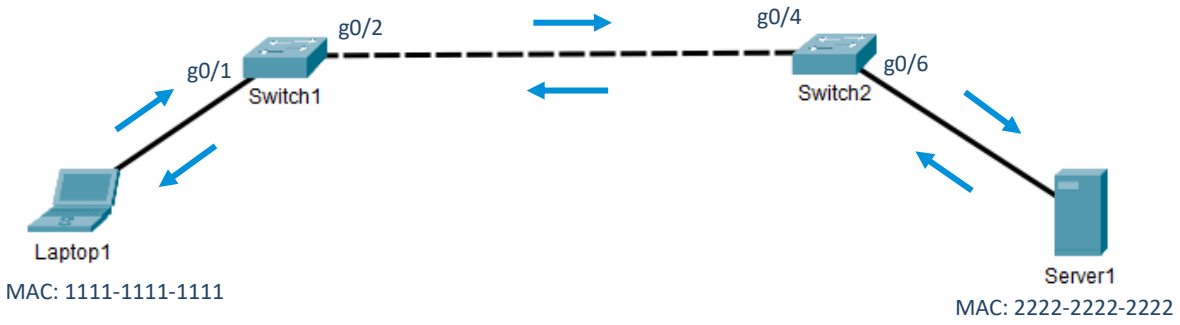
sli.do

#CNA

Spanning Tree Protocol Advanced

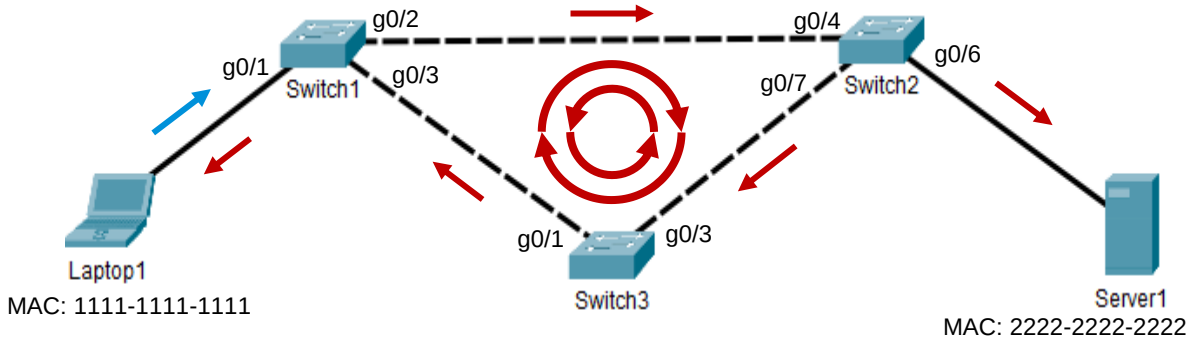
The Problem with Layer 2 Network Redundancy

- Normally, the switches will build and use MAC address tables to help them with the forwarding decisions
- When redundant paths are present, the switches are confused and will forward the received packet everywhere
  - Their MAC address tables become unstable
  - Multiple copies of the same frame (data) are received by all devices, endlessly
  - The links are overloaded
  - This is known as **Layer 2 loop and is very bad situation!**



Switch2 MAC Table

MAC address	Port
1111-1111-1111	g0/4
2222-2222-2222	g0/6

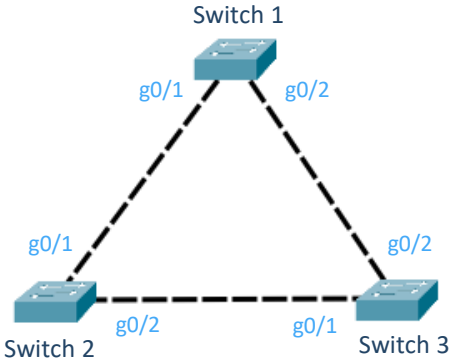


Switch2 MAC Table

MAC address	Port
1111-1111-1111	~~g0/4~~, ~~g0/7~~, ~~g0/4~~…**?**
2222-2222-2222	~~g0/6~~, ~~g0/7~~, ~~g0/4~~…**?**

The STP Algorithm

1. Elect the **Root** switch (a.k.a. Root bridge or just Root)
   - This is the switch with the lowest BID (Bridge ID)
   - BID = Switch Priority and MAC
   - Default priority = 32768
2. Select the **root ports**
   - They have the best (lowest) cost to the Root
   - Selected per switch – maximum one
   - Only the non-Root switches have root ports
3. Select the **designated ports**
   - They have the best (lowest) cost to the Root
   - Selected per segment (connection) – exactly one
4. All other ports go to **blocking state**
   - The role of these ports is called "alternate"



The STP tie-breakers

- If there is a tie situation - the same path cost via different paths, use the following tie-breakers:
  - When selecting Root port or Designated port, chose the neighboring switch which has the lowest Bridge ID
  - If the Bridge ID is the same, select the lowest Port ID (PID)
- Port ID = Port priority and port number

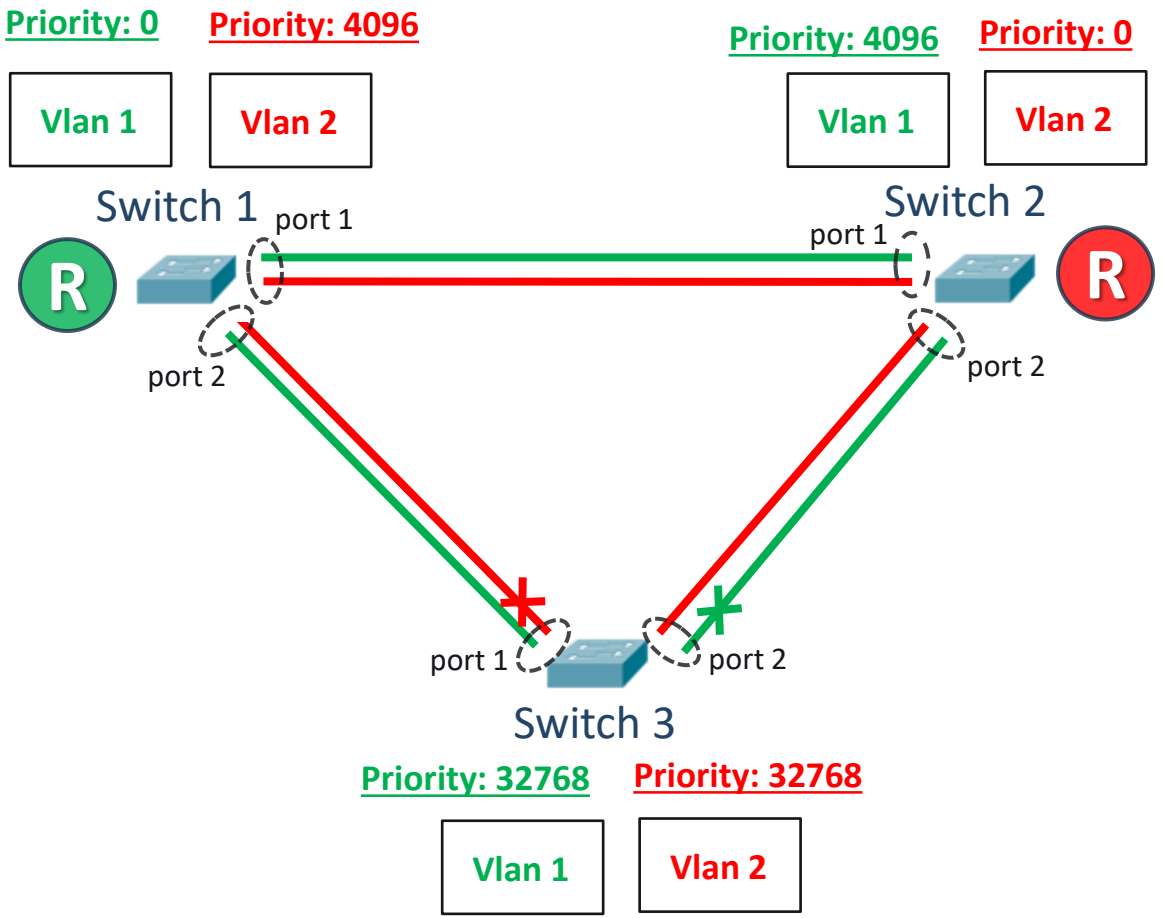
History and Flavors

- **STP** - The original Spanning Tree Protocol, IEEE 802.1D
  - Problem – very slow convergence (between 30 and 50 seconds)
  - Problem – single Root for the entire Layer 2 topology
- **RSTP** - Rapid STP, IEEE 802.1W
  - Much faster convergence (no more timers, introducing the concept of an "edge" port)
  - Still have the issue with single Root for the entire Layer 2 topology
- **PVST+** - Per-VLAN STP, Cisco proprietary
  - Calculates STP for each particular VLAN independently - can have multiple Roots
  - It is also "rapid"
- **MSTP** - Multiple STP, IEEE 802.1S
  - Calculates STP for each "instance" (group of VLANs) – can have multiple Roots
  - It is also "rapid"

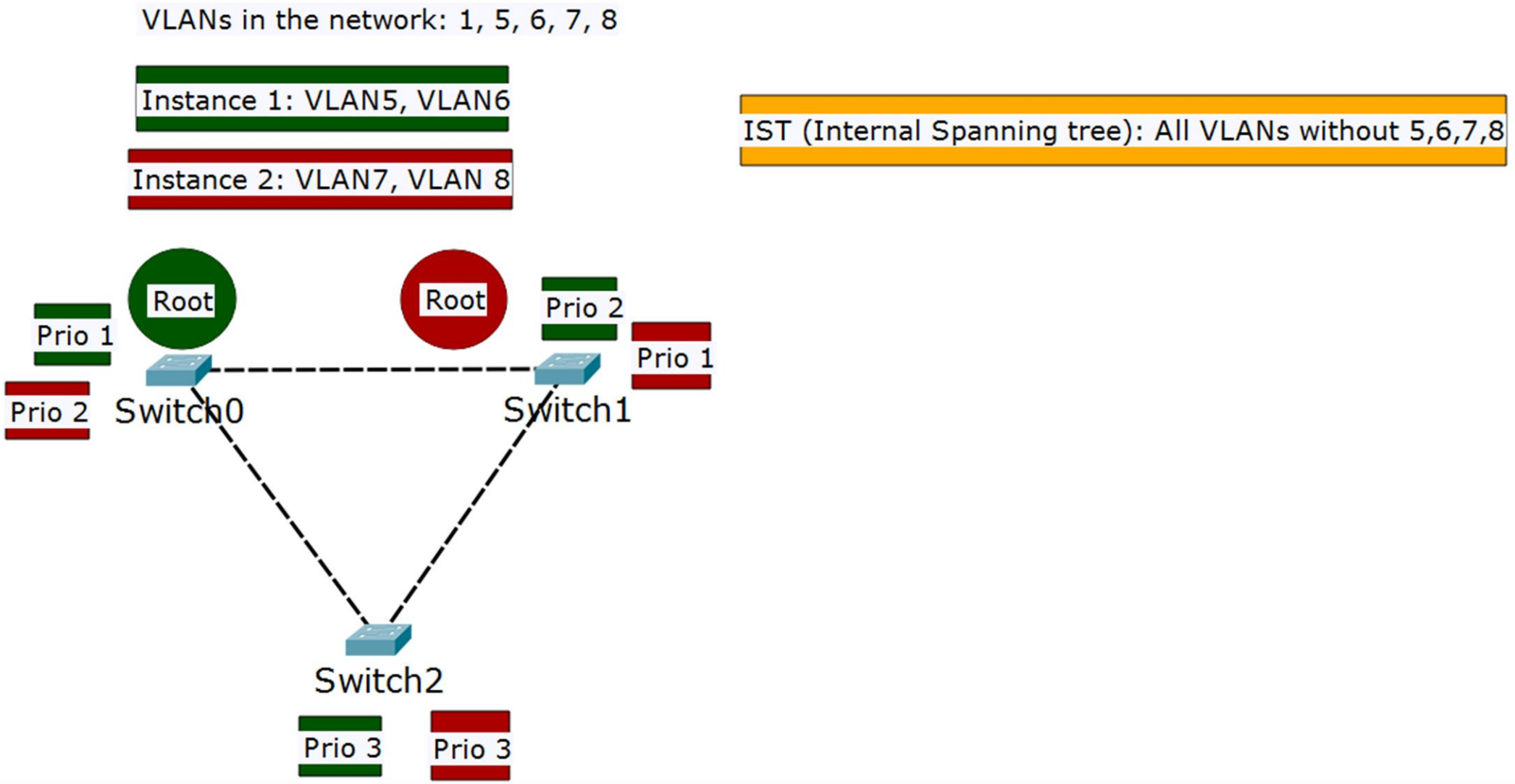
MSTP

- Multiple Spanning Tree Protocol
- Creates multiple instances of the physical STP topology
- Can have **different Root** switches - one for each instance
- Provides **load-sharing** because of the multiple Roots
- It is also "rapid"
- One instance is mapped to one or multiple VLANs
- Needs additional configuration

PVST+: Root Switch Per VLAN



MSTP: Root Switch Per Instance



MSTP Instances

- By default, all VLANs exist inside **Instance 0** (a.k.a. IST)
- The benefit of MSTP comes when:
  - There are multiple VLANs
  - Other (custom) instances are created
  - VLANs are distributed across instances
  - Different priorities are set for the different instances
  - All switches are in the **same region**

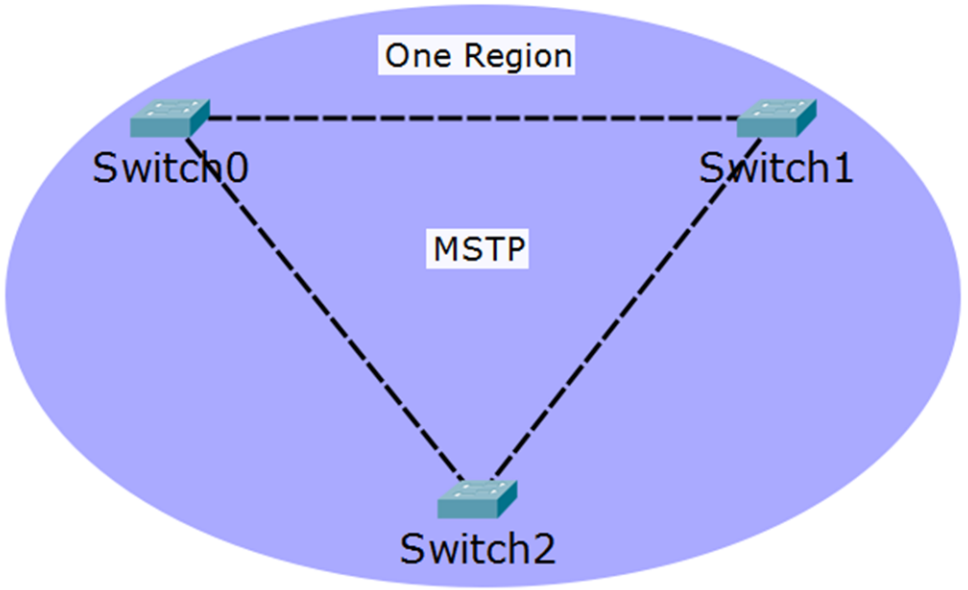
MSTP Configuration

- All switches in the MSTP domain must be in the **same region**
- Three parameters should be configured equally on all switches:
  - Configuration Name (region name)
  - Revision number
  - VLAN-to-instance mappings
- The priorities are configured on a **per instance** basis

MSTP: All Switches in the Same Region

Switch0:

1. Configuration name: SoftUni
2. Revision: 1
3. Instance 1 = vlan 5 and vlan 6
   Instance 2 = vlan 7 and vlan 8



Switch1:

1. Configuration name: SoftUni
2. Revision: 1
3. Instance 1 = vlan 5 and vlan 6
   Instance 2 = vlan 7 and vlan 8

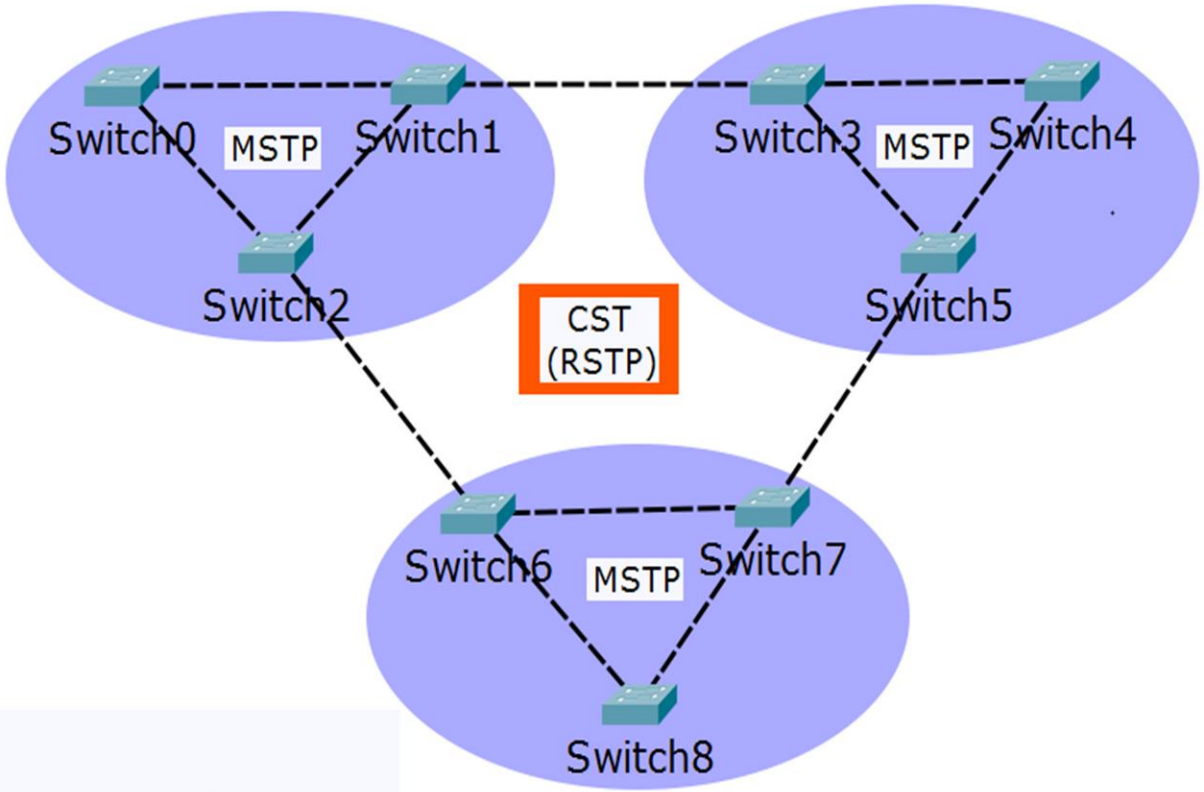
Switch2:

1. Configuration name: SoftUni
2. Revision: 1
3. Instance 1 = vlan 5 and vlan 6
   Instance 2 = vlan 7 and vlan 8

MSTP: Multiple Regions

Region1:

1. Configuration name: SoftUni
2. Revision: 1
3. Instance 1 = vlan 5 and vlan 6
   Instance 2 = vlan 7 and vlan 8



Region2:

1. Configuration name: Region2
2. Revision: 1
3. Instance 1 = vlan 5 and vlan 6
   Instance 2 = vlan 7 and vlan 8

Region3:

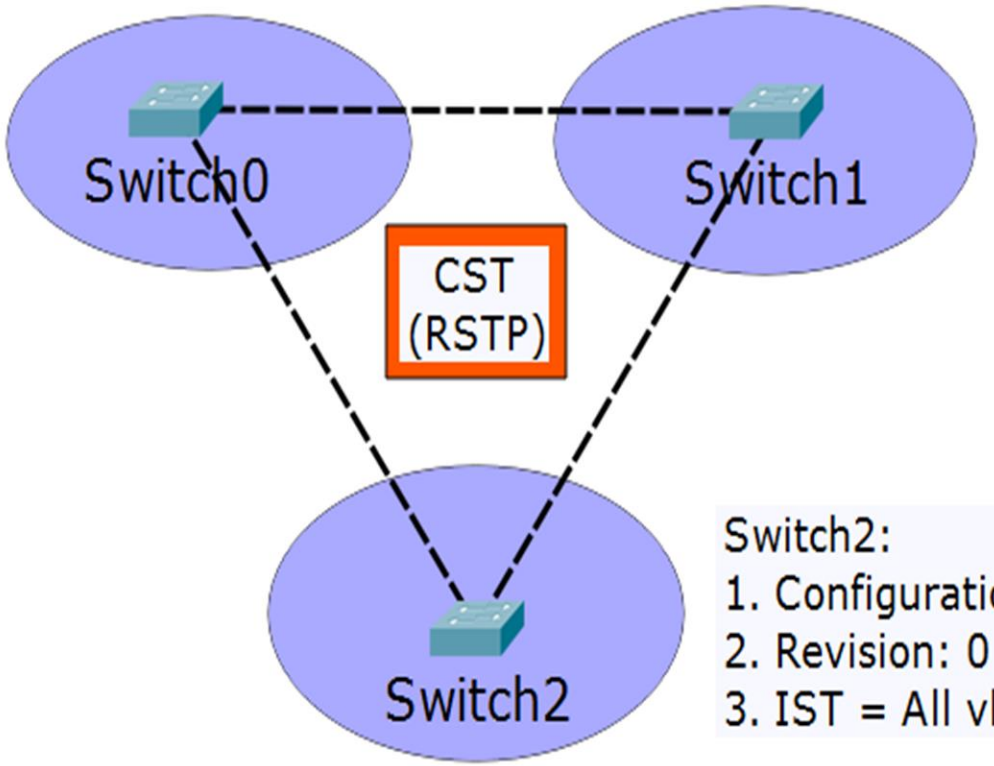
1. Configuration name: Test
2. Revision: 5
3. Instance 1 = vlan 15 and vlan 16
   Instance 2 = vlan 98 and vlan 3

MSTP Defaults

- On many switches, when spanning tree is enabled, they use **MSTP** by default
  …but there is no custom MSTP configuration
- So each switch will be in its own region and CST (**RSTP**) will run between these switches

MSTP Defaults (2)

Switch0:
1. Configuration name: MAC000
2. Revision: 0
3. IST = All vlans



Switch1:
1. Configuration name: MAC111
2. Revision: 0
3. IST = All vlans

Switch2:
1. Configuration name: MAC222
2. Revision: 0
3. IST = All vlans

MSTP vs PVST+

- Two similar protocols, which one is better?
- PVST+ advantages:
  - triggers STP calculation **only if** there is a potential loop in a **particular VLAN**
  - detailed "look" of the network – does not block ports when there is no loop on the trunks for a given VLAN
- PVST+ disadvantages
  - generates **a lot of overhead** in the network
  - proprietary protocol

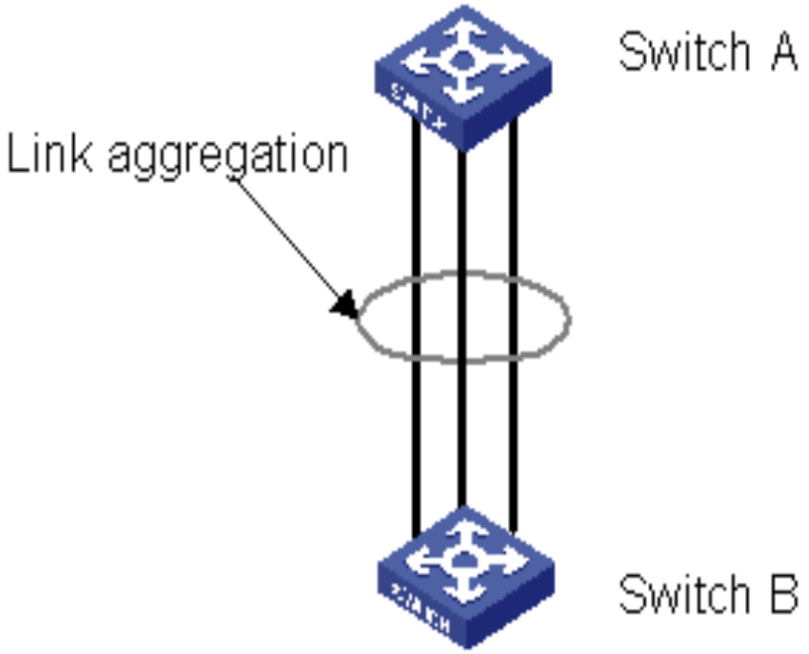
MSTP vs PVST+ (2)

- MSTP advantages:
  - uses **less BPDUs** and generates less overhead
  - open standard
- MSTP disadvantages:
  - not VLAN aware (does not look which VLANs are on the trunk ports)
  - Harder to configure
- Recommendation: use MSTP if you have more than 100 VLANs

Link Aggregation

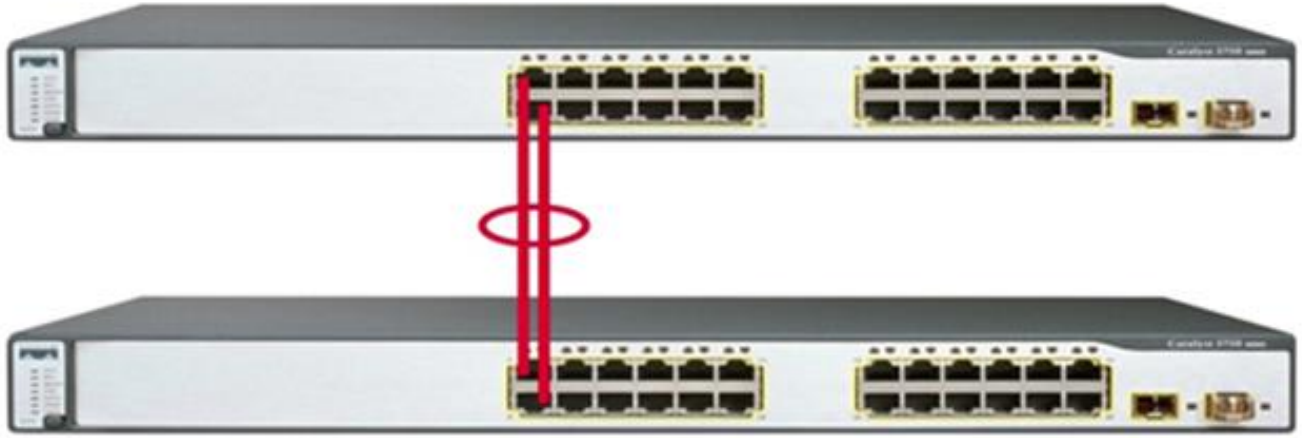
What is Link Aggregation?

- Combination of two or more physical interfaces to create one logical link or port
- Other terms with the same meaning:
  - Port trunking (HPE Provision)
  - EtherChannel (Cisco)
  - Link bundling
  - NIC bonding/teaming
  - etc.



Why to Use Link Aggregation?

- To increase the bandwidth
- To provide redundancy



Link Aggregation Protocols

- Static (no protocol) option for link aggregation
- LACP - IEEE 802.1AX (formerly 802.3ad)
  - Static
  - Dynamic
- PAgP - Cisco Proprietary
- Each protocol has configuration options for different scenarios
- LACP configuration and naming can be interpret (slightly) different between the vendors

Interface Requirements

- In general, the physical interfaces which participate in a link aggregation group, should have the same:
  - Speed
  - Duplex
  - VLAN
  - All other settings – recommended

Load Sharing Modes

Software University

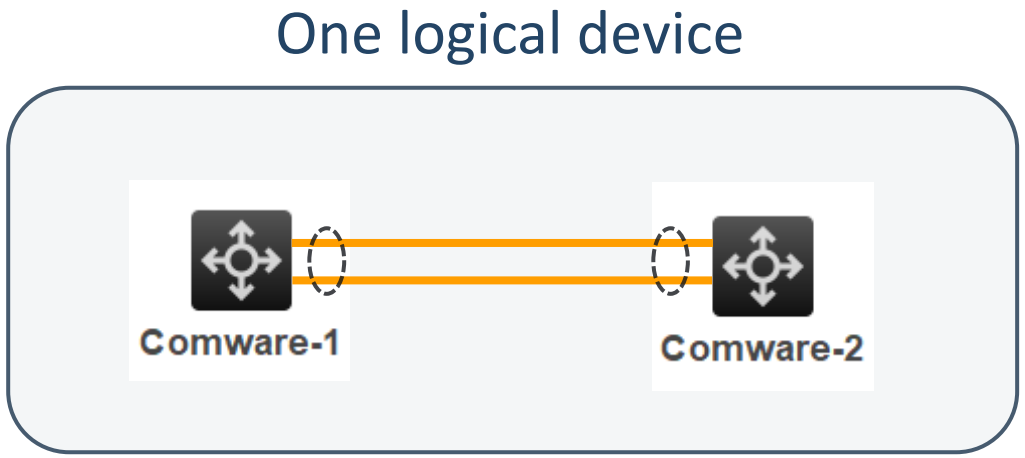
- Multiple physical links form one logical but at the end the traffic uses the physical links
- Which exact link to use - random decision based on "conversations"
- Conversations depend on the load sharing mode

```
Switch(config)#port-channel load-balance ?
  dst-ip        Dst IP Addr
  dst-mac       Dst Mac Addr
  src-dst-ip    Src XOR Dst IP Addr
  src-dst-mac   Src XOR Dst Mac Addr
  src-ip        Src IP Addr
  src-mac       Src Mac Addr
```

Device Stacking: IRF

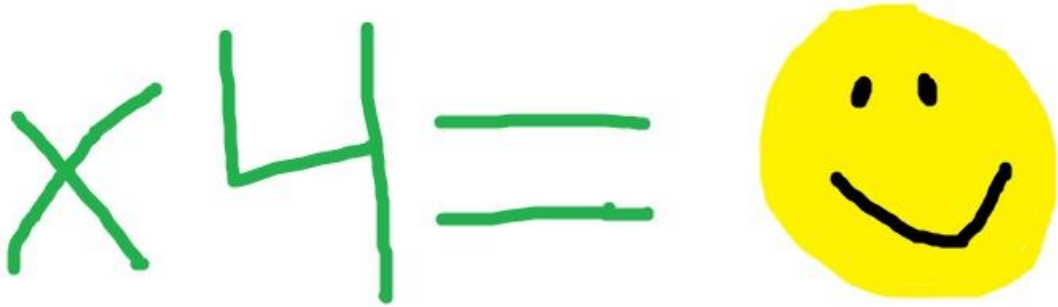
What is Device Stacking?

- Combination of two or more physical devices to form one logical



Why Device Stacking?

- More resilient
- Does not block ports (like STP)
- Simple management
- Simple design



What is IRF?

- IRF – Intelligent Resilient Framework
- HPE proprietary stacking technology
- Uses HPE Comware switches (or routers) only
- The devices need to be from the same type
- One device is elected as a master
- The stack takes the master's configuration

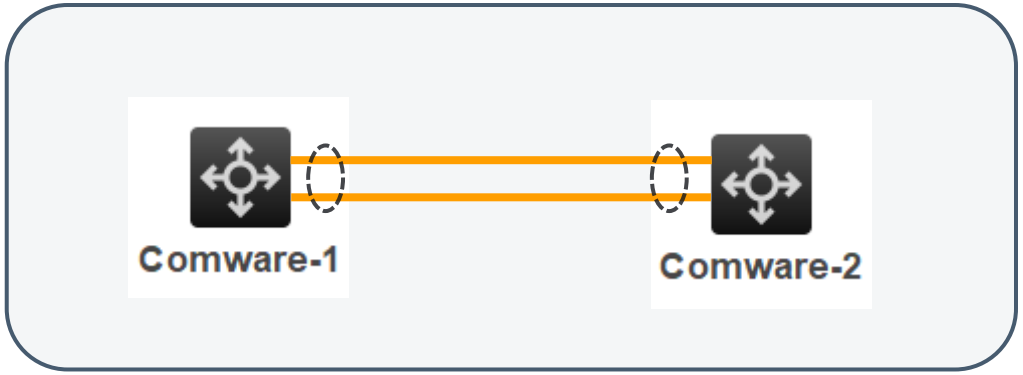
Master Election

- When creating the IRF stack:
  - The member with highest priority wins
  - If no member has a higher priority, the member with the longest system up-time wins (rounded to 10 minutes)
  - If no member has a longer up-time, the member with the lowest bridge MAC address wins
- When joining another device to the IRF stack:
  - The current master wins

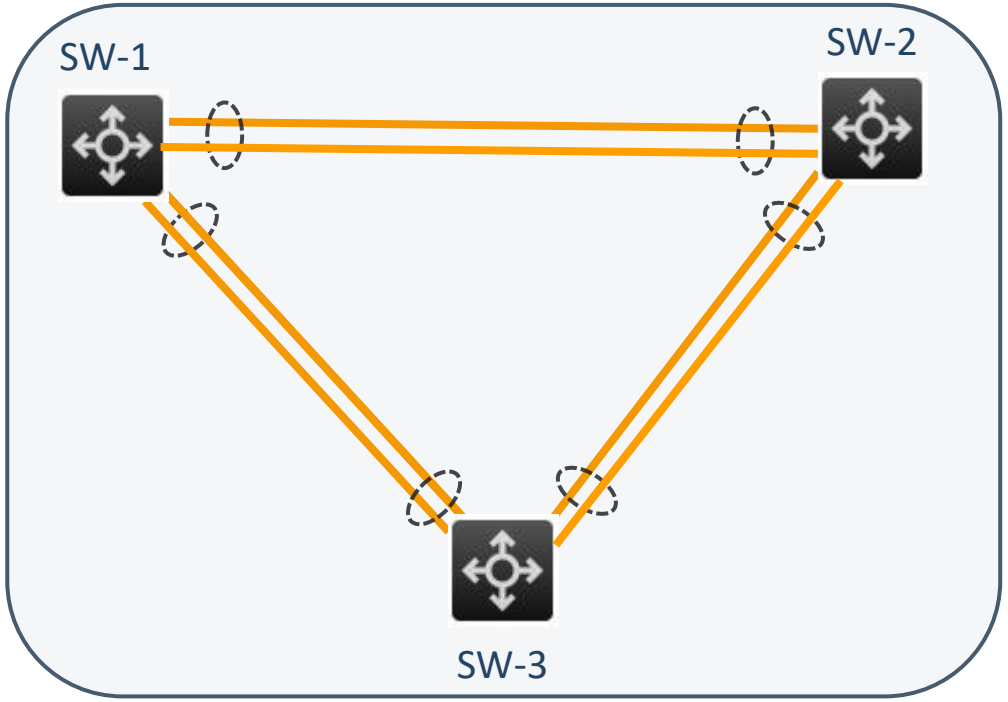
IRF Topologies

Software University

Daisy chain:



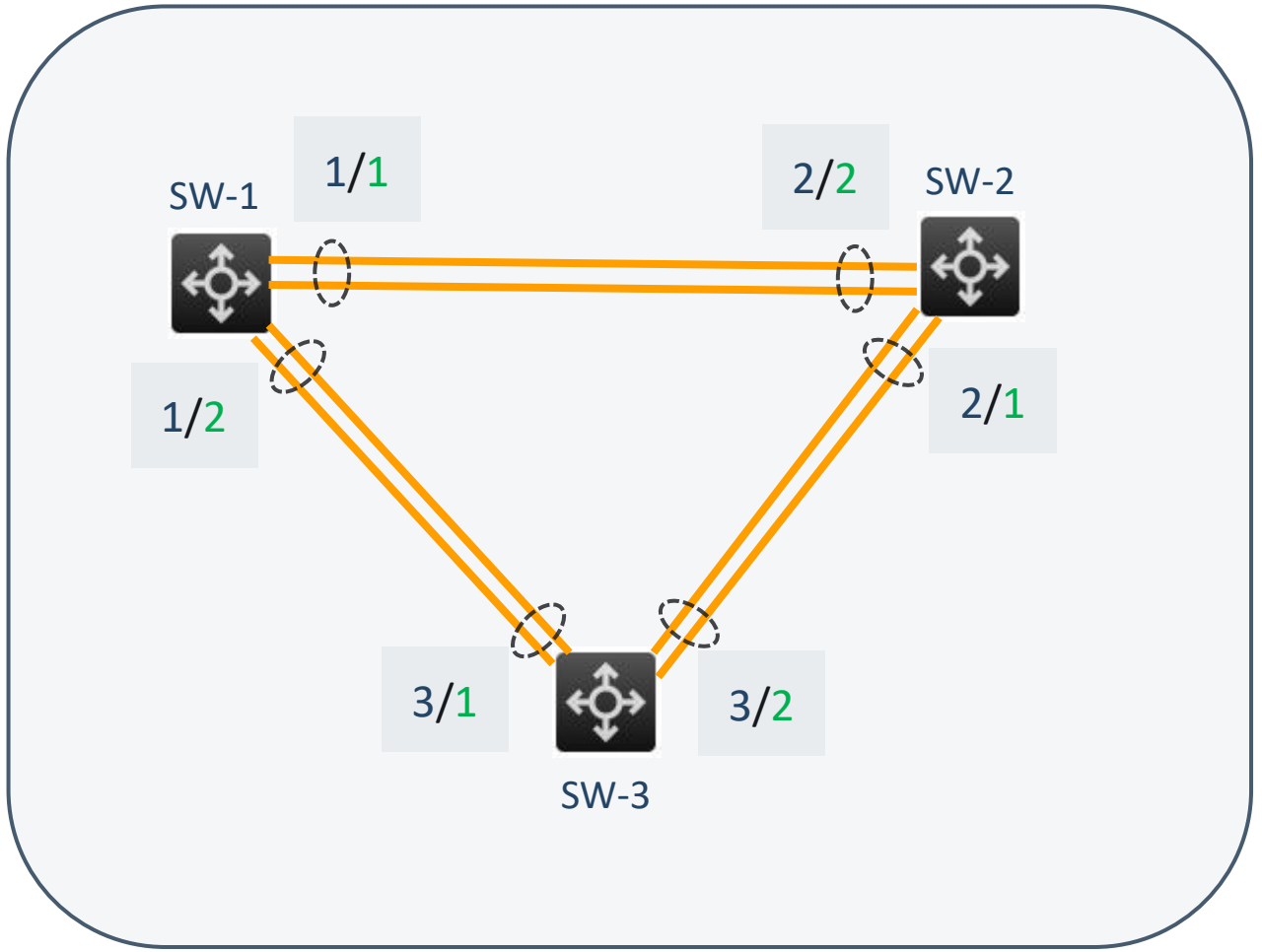
Ring:



IRF Ports and Connections

Single physical cable
Must be at least 10Gbps

x/y IRF port – memberID/port



Only connect IRF port 1 to IRF port 2!

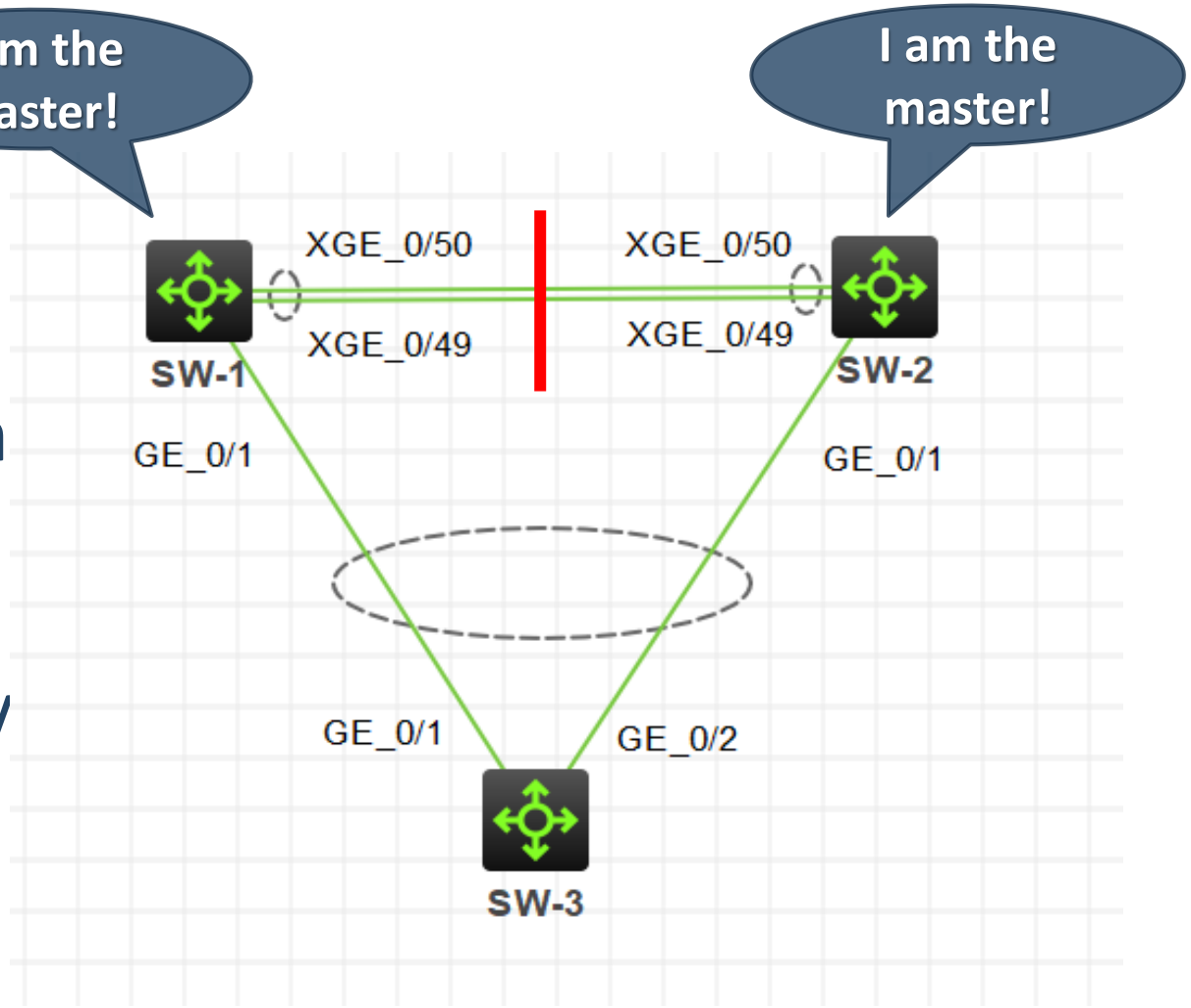
Possible connections	**1/1 -> 2/2**	**1/2 -> 2/1**
Not possible connections	1/1 -> 2/1	1/2 -> 2/2

IRF Split Stack

- When there is a split stack situation, each of the two parts assumes "I am the new master"
- This can create problems
- Additional Multi Active Detection (MAD) algorithm needs to be configured:
  - LACP MAD
  - BFD MAD
  - ARP MAD

Detect IRF Split Stack with LACP MAD

- Additional Comware switch is required
- Extended LACPDU are exchanged
- If split stack, MAD triggers election
- Smaller member ID wins
- The other device(s) put in recovery

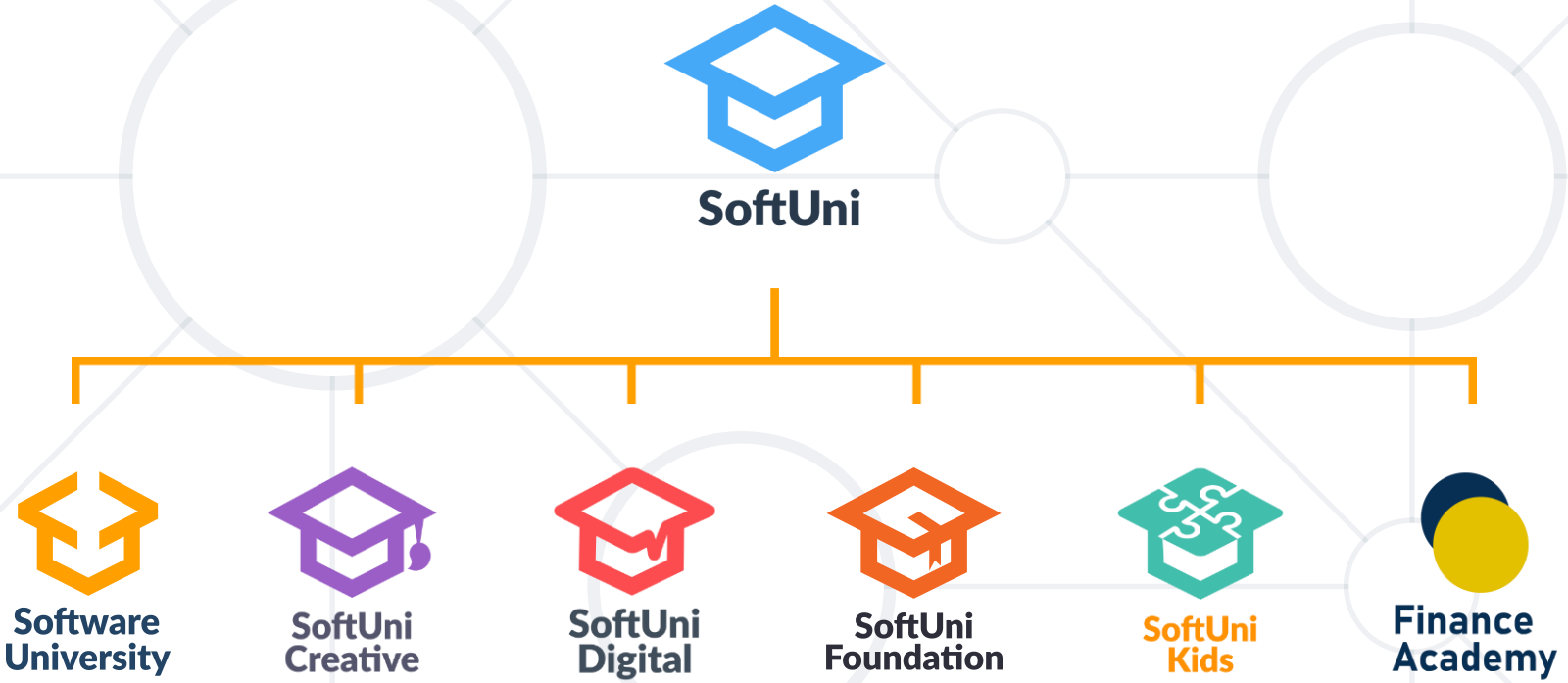


Demonstration

Summary

1. Spanning tree protocol advanced:
   - MSTP and PVST+
2. Link aggregation
3. Device stacking technologies - IRF

Questions?

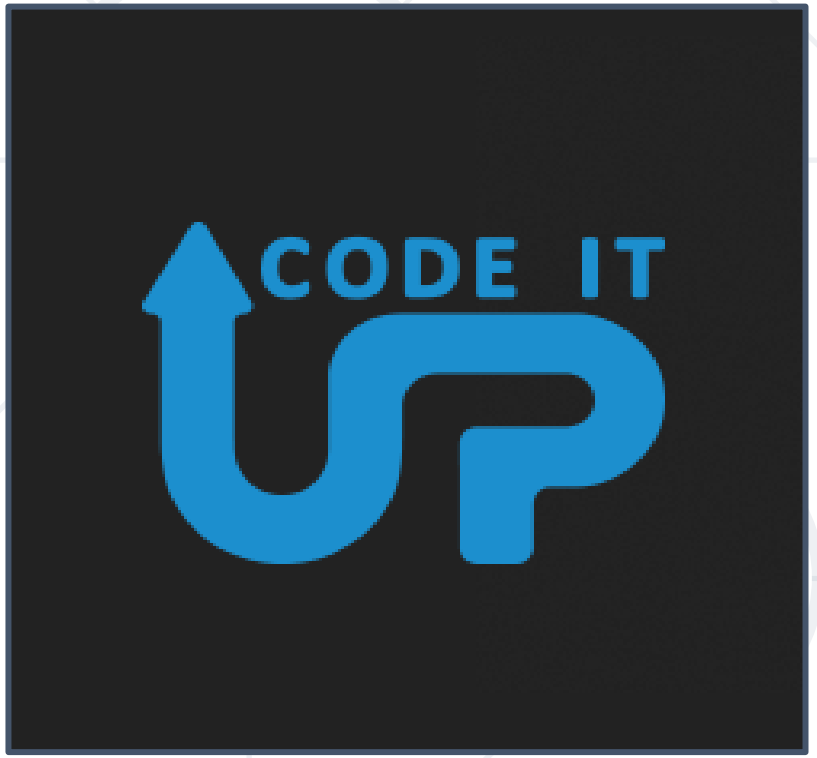


SoftUni Diamond Partners

Software University

Educational Partners

Software University



Trainings @ Software University (SoftUni)

- Software University – High-Quality Education, Profession and Job for Software Developers
  - softuni.bg, about.softuni.bg
- Software University Foundation
  - softuni.foundation
- Software University @ Facebook
  - facebook.com/SoftwareUniversity
- Software University Forums
  - forum.softuni.bg