# Other Layer-2 Features

## Network Infrastructure Workshop

Last updated October 2019

SOMNOG 6
Somali Network Operators Group
WORKSHOPS AND CONFERENCE
21 - 26 October 2023

This document is a result of work by the Network Startup Resource Center (NSRC at http://www.nsrc.org). This document may be freely copied, modified, and otherwise re-used on the condition that any re-use acknowledge the NSRC as the original source.

# Other Layer-2 Features

- Link Aggregation
- Rapid Spanning Tree
- Multiple Spanning Tree
- Switch Configuration Advice:
  - Network Management
  - Documentation

# Link Aggregation

- Also known as *port bundling, link bundling*
- You can use multiple links in parallel as a single, logical link
  - For increased capacity
  - For redundancy (fault tolerance)
- LACP (Link Aggregation Control Protocol) is a standardized method of negotiating these bundled links between switches
- Proprietary methods exist too (PAgP, EtherChannel)

# LACP Operation

- Two switches connected via multiple links will send LACPDU packets, identifying themselves and the port capabilities
- They will then automatically build the logical aggregated links, and then pass traffic.
- Switch ports can be configured as active or passive

# LACP Operation
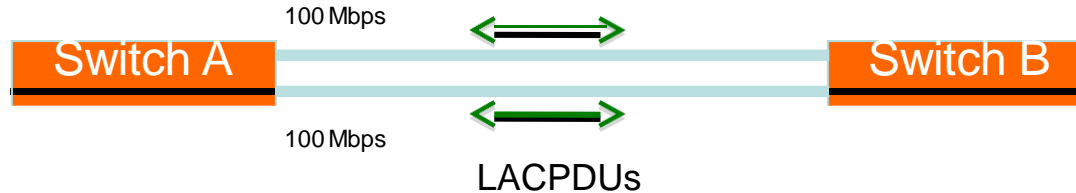


- Switches A and B are connected to each other using two sets of Fast Ethernet ports

- LACP is enabled and the ports are turned on

- Switches start sending LACPDUs, then negotiate how to set up the aggregation

# LACP Operation

Switch A  100 Mbps  Switch B

100 Mbps

200 Mbps logical link
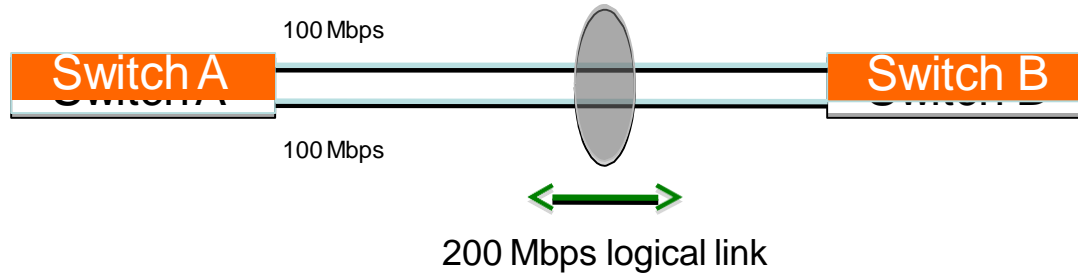
- The result is an aggregated 200 Mbps logical link
- The link is also fault tolerant: If one of the member links fail, LACP will automatically take that link off the bundle, and keep sending traffic over the remaining link

# Distributing Traffic in Bundled Links

- Bundled links distribute frames using a hashing algorithm, based on:

  - Source and/or Destination MAC address
  - Source and/or Destination IP address
  - Source and/or Destination Port numbers

- This can lead to unbalanced use of the links, depending on the nature of the traffic

- Always choose the load-balancing method that provides the most distribution

# Questions?

# Rapid Spanning Tree (802.1w)

- Backwards-compatible with 802.1d
- Convergence is **much** faster
  - Communication between switches is more interactive
- Edge ports don't participate
  - Edge ports transition to forwarding state immediately
  - If BPDUs are received on an edge port, it becomes a non-edge port to prevent loops
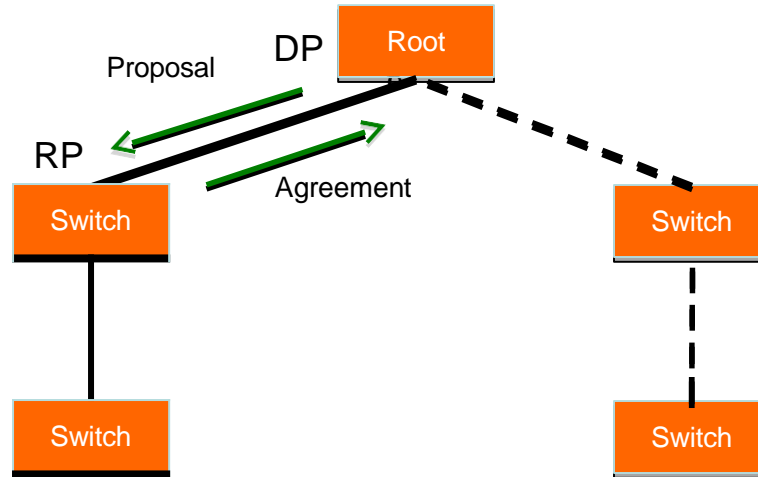
# Rapid Spanning Tree (802.1w)

- Defines these port roles:
  - Root Port (same as with 802.1d)
  - Alternate Port
    - A port with an alternate path to the root
  - Designated Port (same as with 802.1d)
  - Backup Port
    - A backup/redundant path to a segment where another bridge port already connects.

# Rapid Spanning Tree (802.1w)

- Synchronization process uses a handshake method
  - After a root is elected, the topology is built in cascade, where each switch proposes to be the designated bridge for each point-to-point link
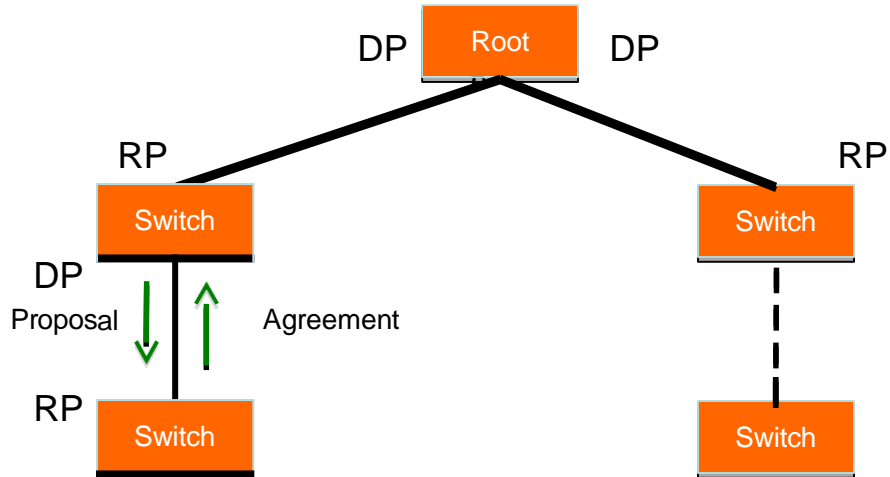  - While this happens, all the downstream switch links are blocking
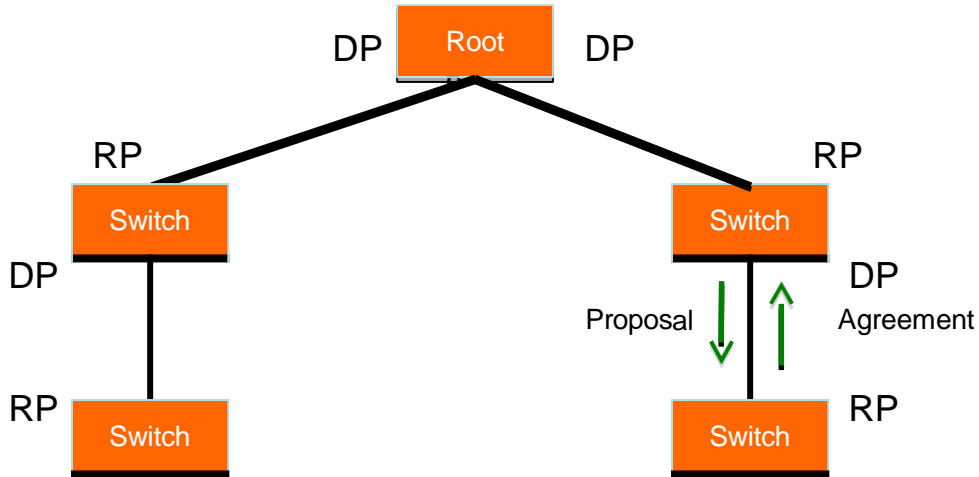
# Rapid Spanning Tree (802.1w)

# Rapid Spanning Tree (802.1w)
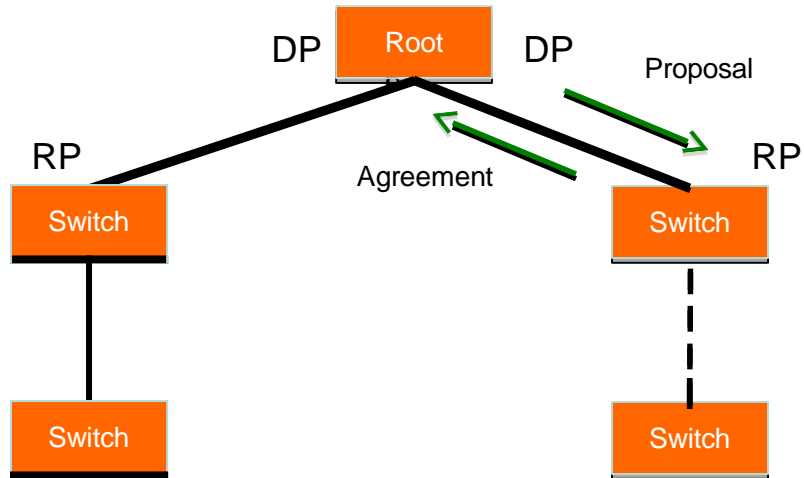
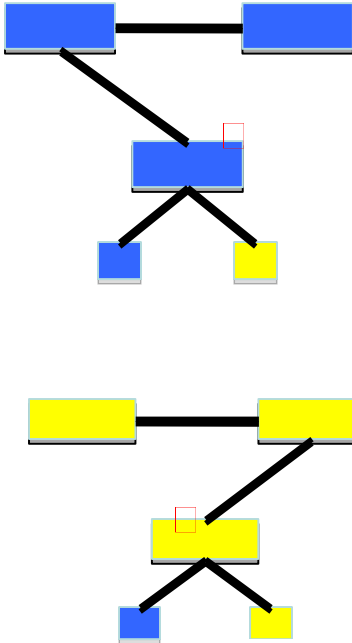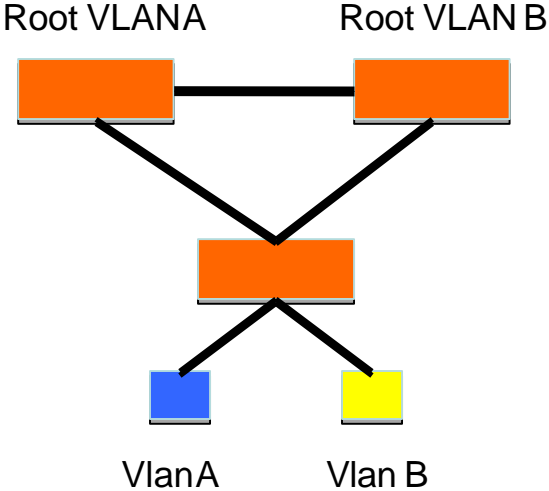# Rapid Spanning Tree (802.1w)

# Rapid Spanning Tree (802.1w)

# Questions?

# Multiple Spanning Tree (802.1s)

- Backwards-compatible with 802.1d
- Compatible with RSTP
- Allows separate spanning trees per VLAN group
  - Different topologies allow for load balancing between links
  - One or more VLANs are assigned (mapped) to an "instance" of MST (MSTI)
  - A particular VLAN can be mapped to only one MSTI

# Multiple Spanning Tree (802.1s)

Root VLANA          Root VLAN B

VlanA          Vlan B

# Multiple Spanning Tree (802.1s)

- MST Instances
  - Groups of VLANs are mapped to particular Spanning Tree instances
  - These instances will represent the alternative topologies, or forwarding paths
  - You specify a root and alternate root for each instance

# Multiple Spanning Tree (802.1s)

- MST Region
  - Switches are members of a region if they have the same set of attributes:
    - MST configuration name
    - MST configuration revision
    - Instance-to-VLAN mapping
  - A digest of these attributes is sent inside the BPDUs for fast comparison by the switches
  - One region is usually sufficient

# Multiple Spanning Tree (802.1s)

- CST = Common Spanning Tree
  - Defined in 802.1q standard
  - In order to interoperate with other versions of Spanning Tree, MST needs a common tree that contains all the islands, including other MST regions
  - One spanning-tree instance for the entire bridged network regardless of the number of VLANs or regions

# Multiple Spanning Tree (802.1s)

- IST = Internal Spanning Tree
  - Internal to the Region, that is
  - Within each MST region, the MSTP maintains multiple spanning-tree instances
  - Instance 0 is a special instance known as IST, which "extends" CST inside the MST region
  - MSTI instance 0 is always present if the switch is running MSTP
  - Presents the entire region as a single virtual bridge to the CST outside

# Multiple Spanning Tree (802.1s)
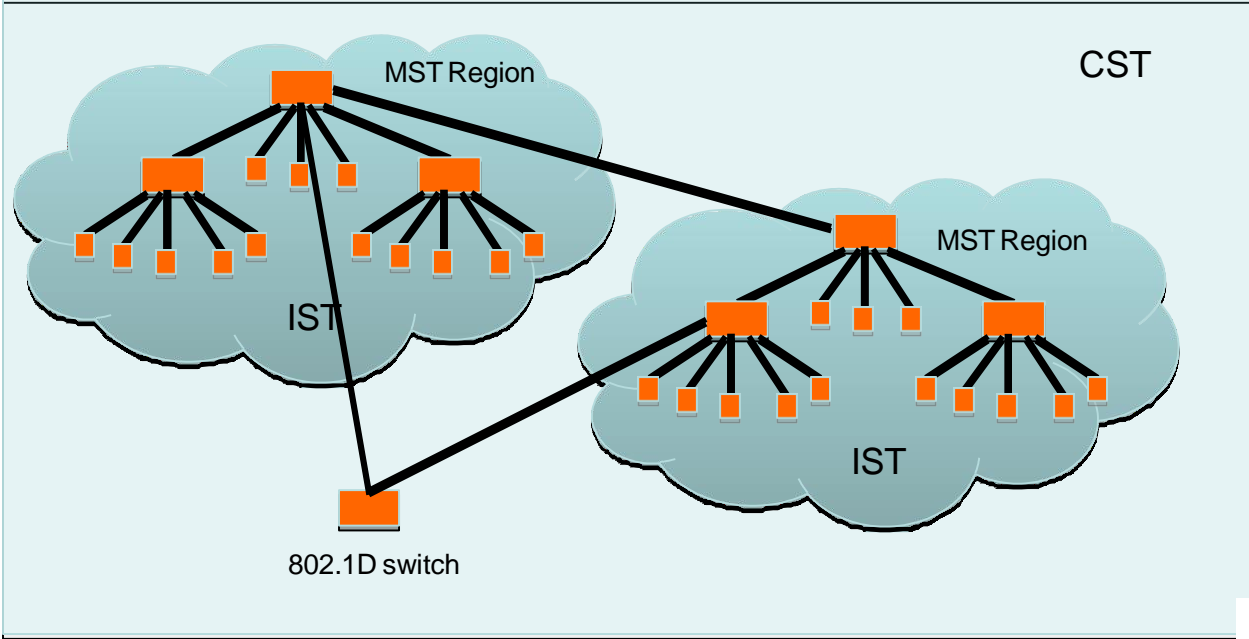
# Multiple Spanning Tree (802.1s)

- Design Guidelines
  - Determine relevant forwarding paths, and distribute your VLANs equally into instances matching these topologies
  - Assign different root and alternate root switches to each instance
  - Make sure all switches match region attributes
  - Do not assign VLANs to instance 0, as this is used by the IST

# Questions?

# Network Management

- Enable SNMP traps and/or syslog
  - Collect and process in centralized log server
    - Spanning Tree Changes
    - Duplex mismatches
    - Wiring problems
- Monitor configurations
  - Use RANCID or Oxidized to report any changes in the switch configuration

# Network Management

- Collect forwarding tables with SNMP
  - Allows you to find a MAC address in your network quickly
  - You can use simple text files + grep, or a web tool with DB backend
- Enable LLDP (or CDP or similar)
  - Shows how switches are connected to each other and to other network devices

# Documentation

- Document where your switches are located
  - Name switch after building name
    - E.g. building1-sw1
  - Keep files with physical location
    - Floor, closet number, etc.

- Document your edge port connections
  - Room number, jack number, server name

# Questions?