Taller de Diseño de Redes de Campus

Diseño de capa 2

UNIVERSITY OF OREGON

Conceptos de capa 2

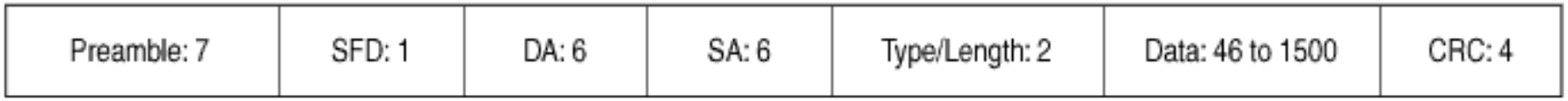
- Los protocolos de capa 2 controlan el acceso a un medio común (cobre, fiber, ondas electromagnéticas)
- Ethernet es el estándar de-facto hoy día
  - Razones:
    - Simple
    - Barato
    - Los fabricantes continúan haciéndolo más rápido

Funciones de Ethernet

- Identificación de la fuente y el destino
  - Direcciones MAC
- Detectar y evitar colisiones
  - Escuchar y esperar a que el canal esté libre
  - Si una colisión ocurre, esperar un tiempo aleatorio antes de reintentar
    - Esto se conoce como CSMA-CD: Carrier Sense Multiple Access with Collision Detection

Trama Ethernet

Normal Ethernet frame



- SFD = Start of Frame Delimiter
- DA = Destination Address
- SA = Source Address
- CRC = Cyclick Redundancy Check

Evolución de Topologías Ethernet

- Bus
  - Todos en el mismo cable coaxial
- Estrella
  - Un dispositivo central conecta a todos los nodos
    - Primero con hubs (tráfico repetido)
    - Luego con switches (tráfico “puenteado”)
  - Se estandarizan los modelos de cableado estructurado

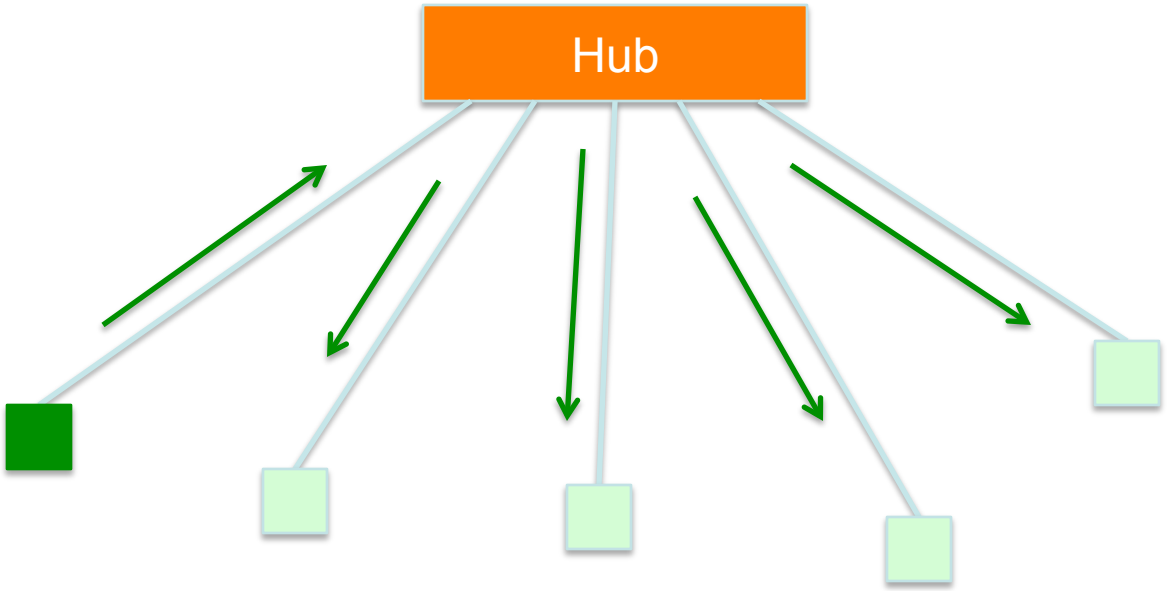
Beneficios de la topología de estrella

- Es modular:
  - Cables independientes para cada nodo
  - Tráfico independiente en cada cable
  - Se puede agregar una segunda capa de switches para repetir lo anterior
  - Siempre diseñe pensando en modularidad

Hub

- Recibe una trama en un puerto y la repite en todos los demás puertos.
- El dominio de colisión abarca todo el hub
- El tráfico termina yendo a sitios donde no es necesario

Hub



Cada trama enviada llega a todos los demás nodos.
Los hubs también se llaman “repetidores” porque repiten todo lo que escuchan

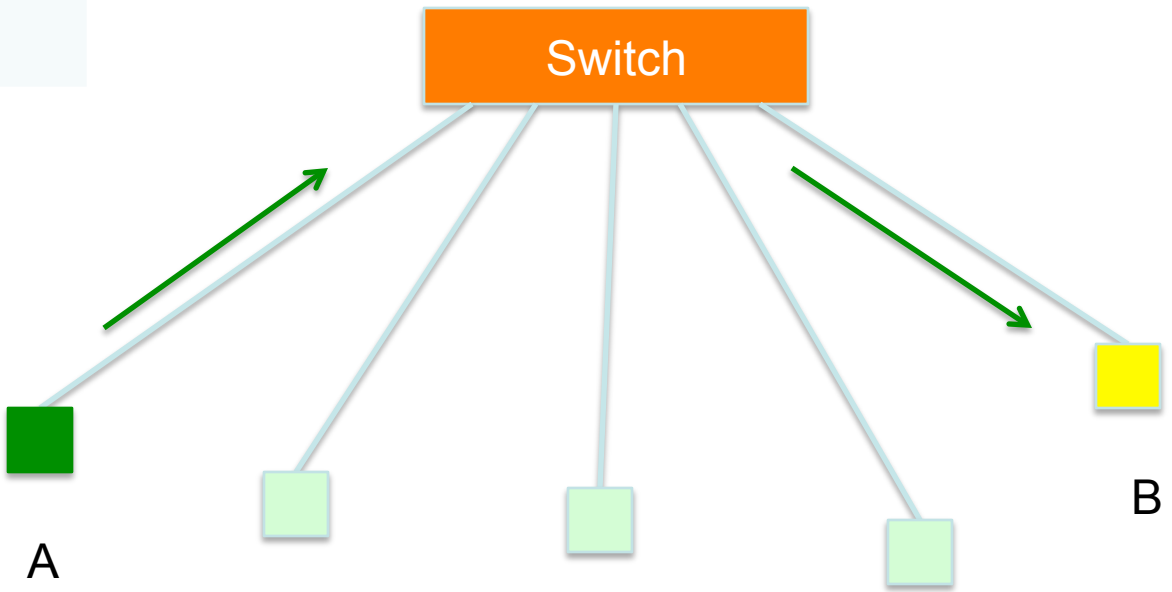
Switch

- ***Aprende*** la ubicación de cada nodo mirando la dirección origen de cada trama, y construye una **tabla de reenvío**
- ***Reenvía*** cada trama sólo a través del puerto donde se encuentra el receptor
  - Reduce el dominio de colisión
  - Utiliza el cable más eficientemente
  - Los nodos no pierden tiempo verificando tramas que no les pertenecen

Switch

Tabla de reenvío

Dirección	Puerto
AAAAAAAAAAAA	1
BBBBBBBBBBBB	5



Switches y Broadcast

- Un switch aún tiene que hacer broadcast con algunas tramas:
  - Cuando el destino no se encuentra en la tabla
  - Cuando el destino de la trama es la dirección broadcast (FF:FF:FF:FF:FF:FF)
  - Cuando el destino de la trama es una dirección multicast
- Así que… los switches no reducen el dominio de broadcast!

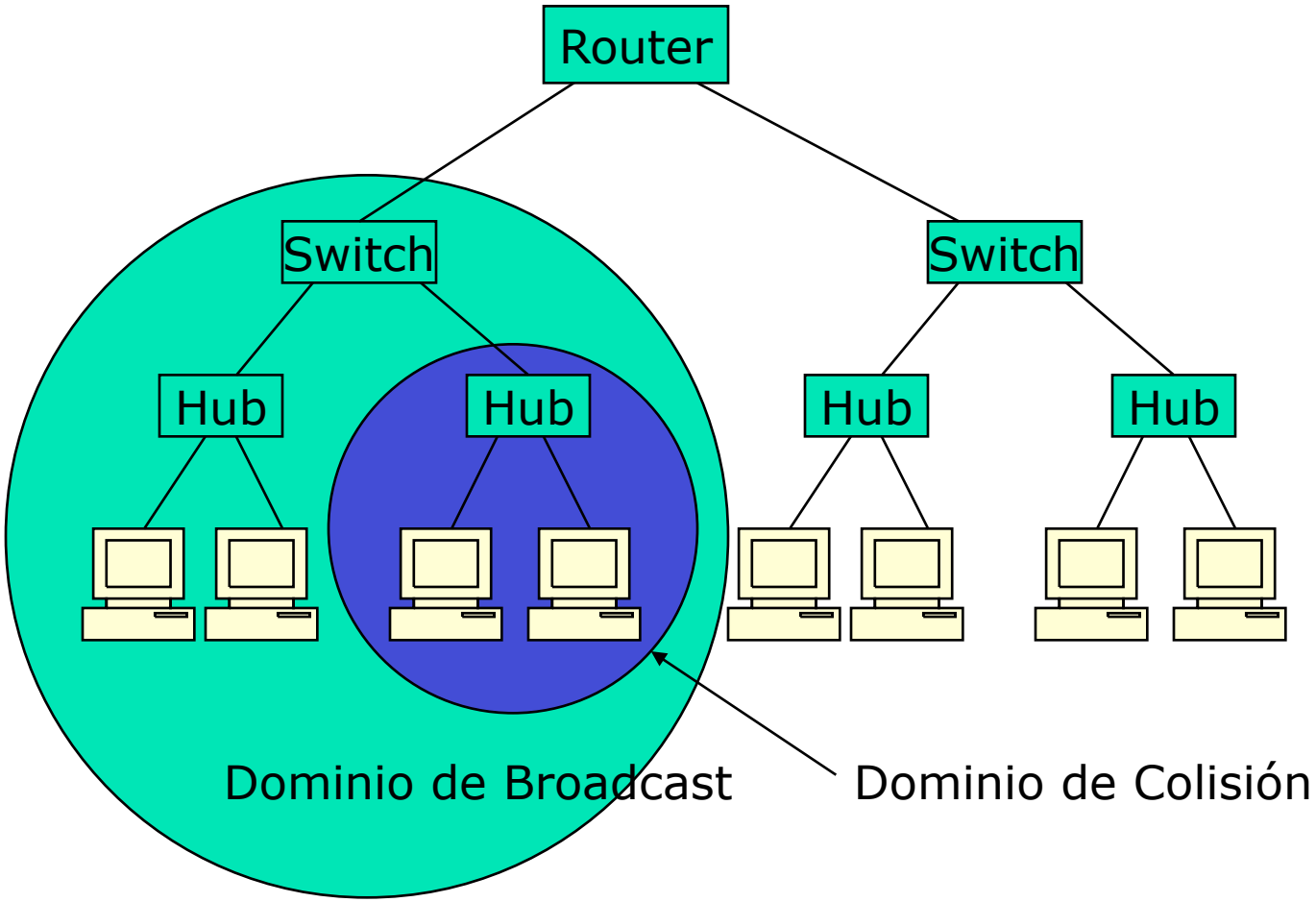
Switch vs. Router

- Los routers más o menos hacen con los paquetes IP lo que los switches hacen con las tramas ethernet
  - Un enrutador mira la dirección destino del paquete IP y la busca en su **tabla de enrutamiento**
- Algunas diferencias:
  - Los paquetes IP viajan dentro de las tramas ethernet
  - Las redes IP se pueden segmentar en subredes
  - Los switches en general no saben nada de IP, sólo de tramas ethernet

Switch vs. Router

- Los enrutadores no reenvían los broadcasts ethernet, así que:
  - Los switches reducen el dominio de colisión
  - Los enrutadores reducen el dominio de broadcast
- Esto cobra suma importancia cuando se intenta diseñar redes jerárquicas que puedan crecer de forma sostenible

Dominios de Tráfico



Dominios de Tráfico

- Intente eliminar los dominios de colisión
  - Desházase de los hubs!
- Intente mantener su dominio de broadcast dentro del un umbral de 250 máquinas conectadas simultáneamente
  - Segmente su red utilizando enrutadores

Pautas de diseño de redes capa 2

- Siempre conecte jerárquicamente
  - Si hay múltiples switches en un edificio, designe uno de ellos como switch de agregación
  - Ubique el switch de agregación cerca del punto de entrada al edificio (panel de fibra)
  - Ubique los switches de acceso cerca de los usuarios (ej. uno por piso)
    - Recuerde que la longitud máxima para Cat5 es 100 metros

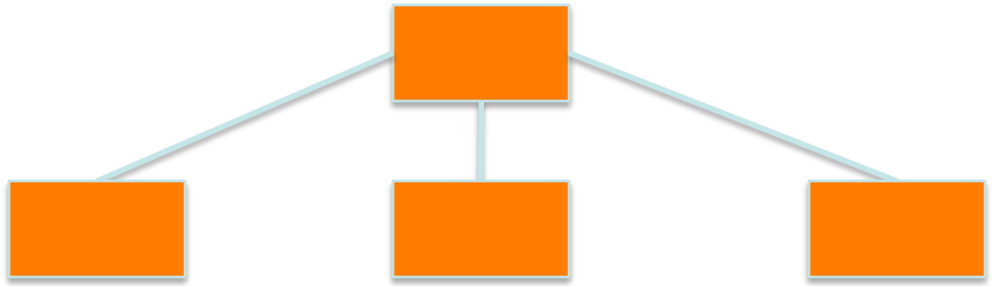
Edificios y subredes

- Es común encontrar correspondencia entre edificios y subredes
  - Switching dentro del edifcio
  - Routing entre edificios
- Esto dependerá del tamaño de la red
  - Edificios con pocas máquinas pueden compartir una subred
  - Edifcios con gran número de máquinas pueden tener distintas subredes (por ej. una subred en cada nivel)

Red de Edificio

Minimice el camino entre elementos

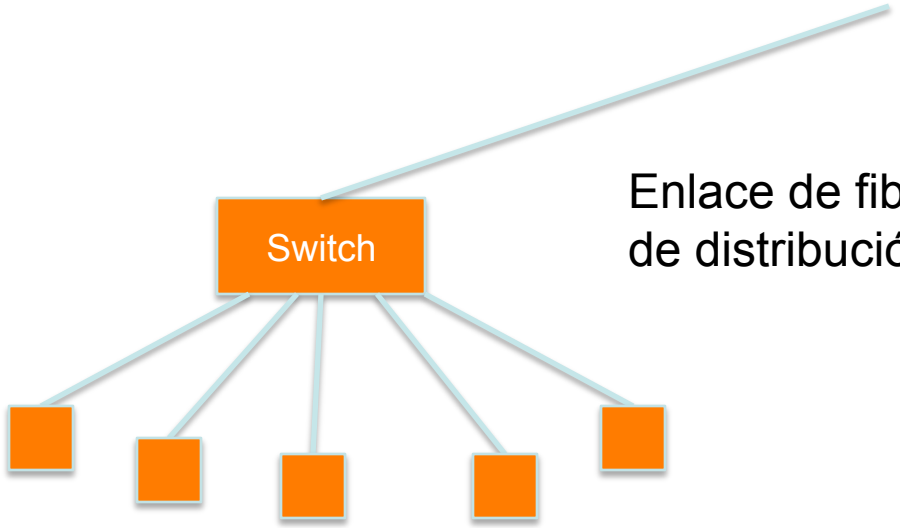
✗



✔

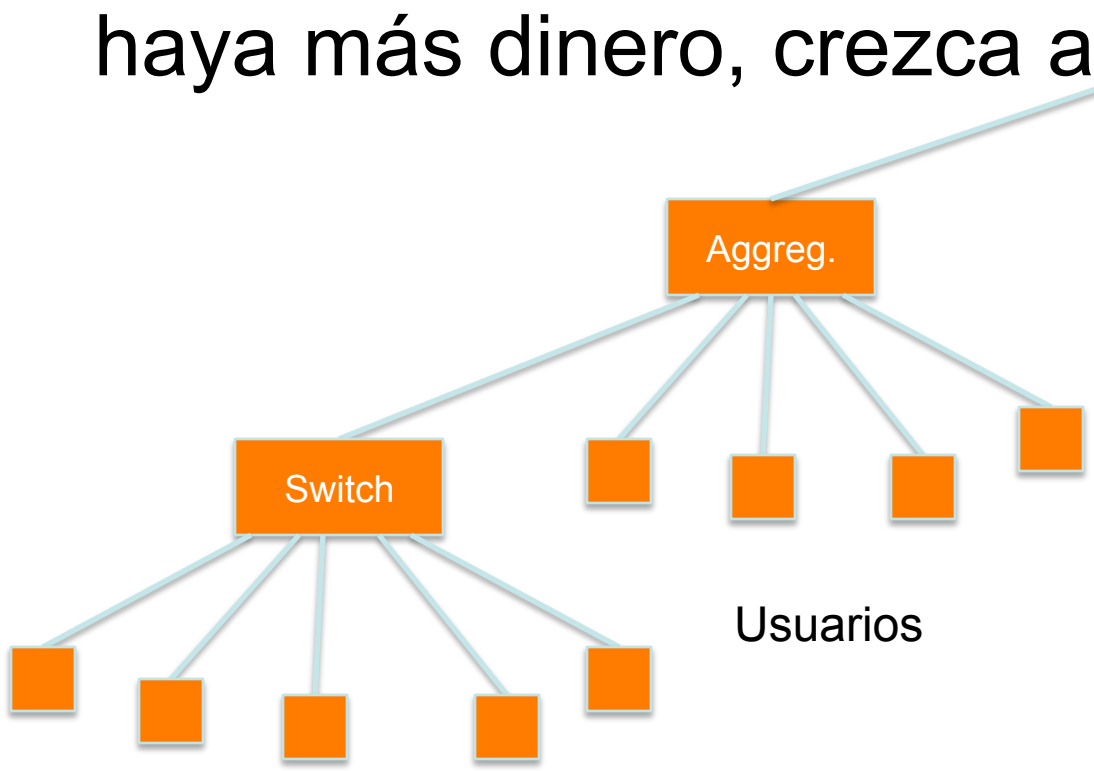
Incremente poco a poco

- Empiece con algo pequeño



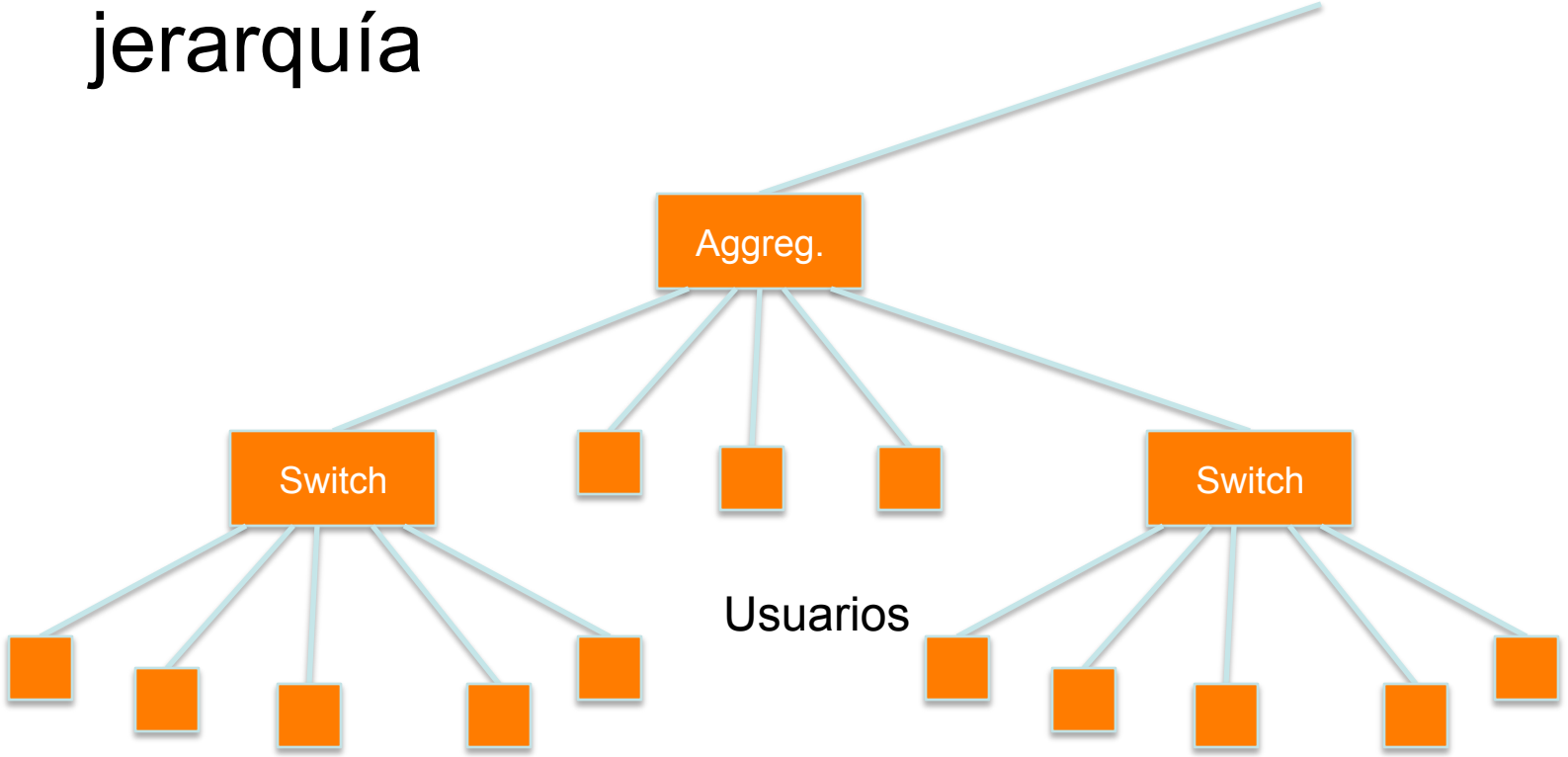
Incremente poco a poco

- A medida que la demanda aumente y haya más dinero, crezca así:



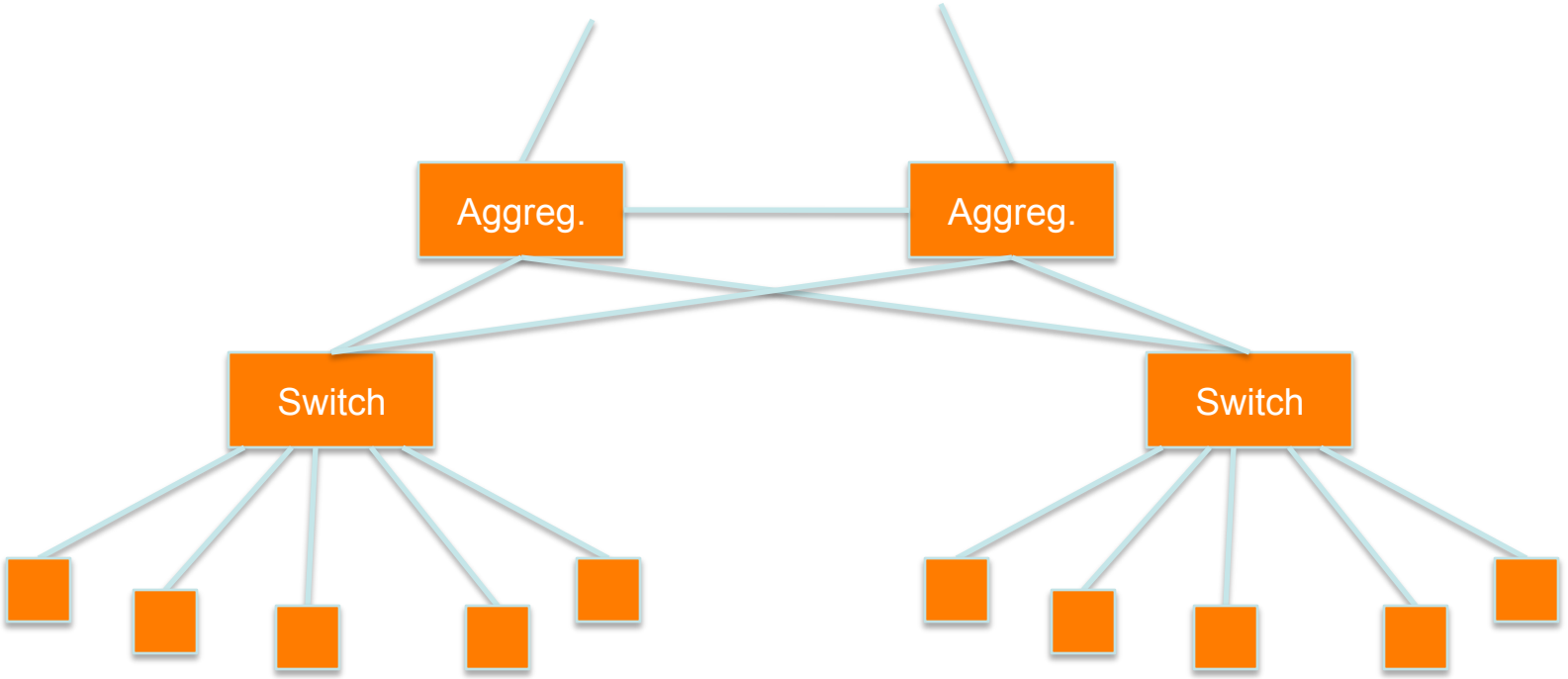
Incremente poco a poco

- Y siga creciendo dentro de la misma jerarquía



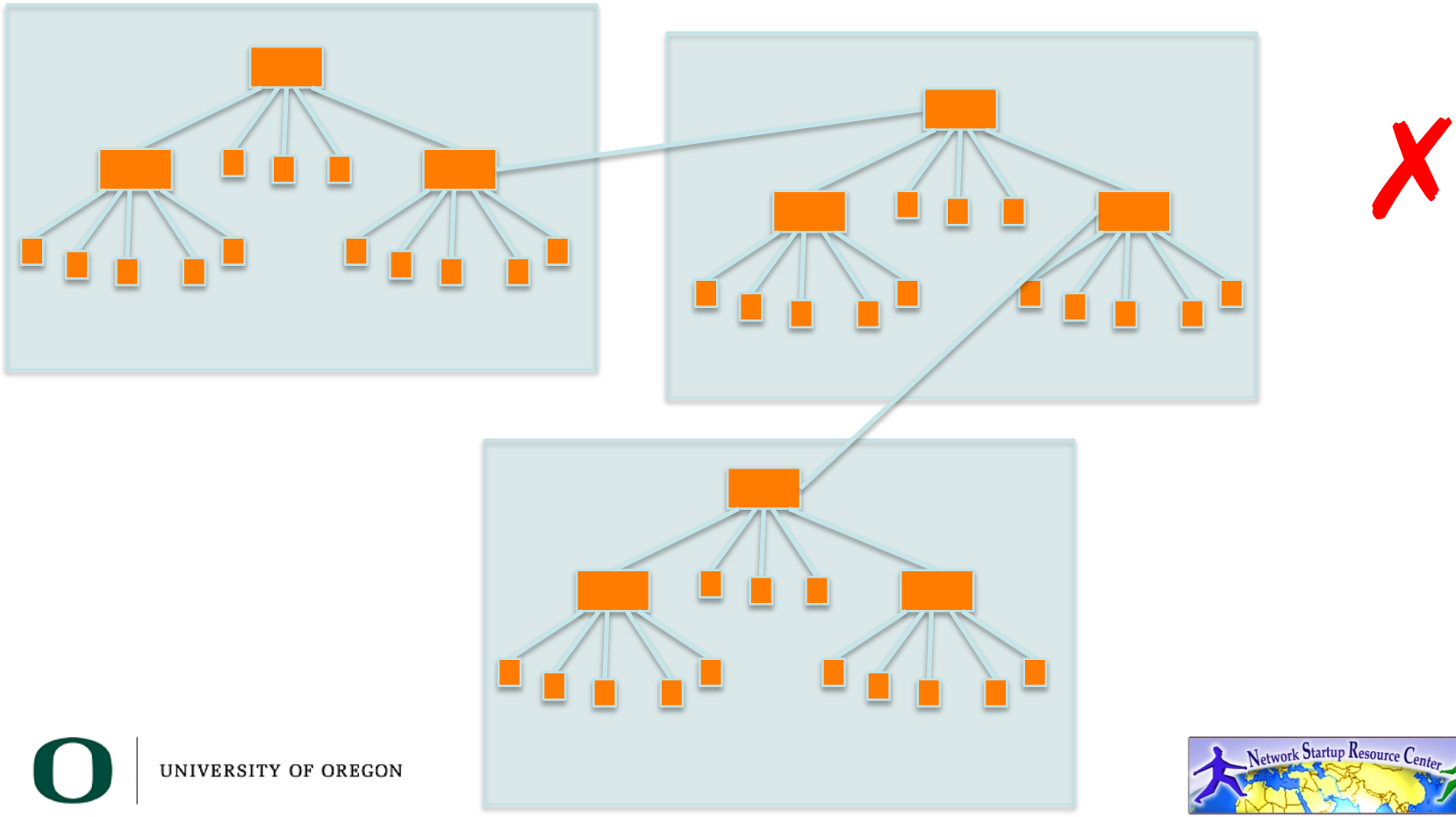
Incremente poco a poco

- En este punto, puede agregar otro switch dorsal redundante

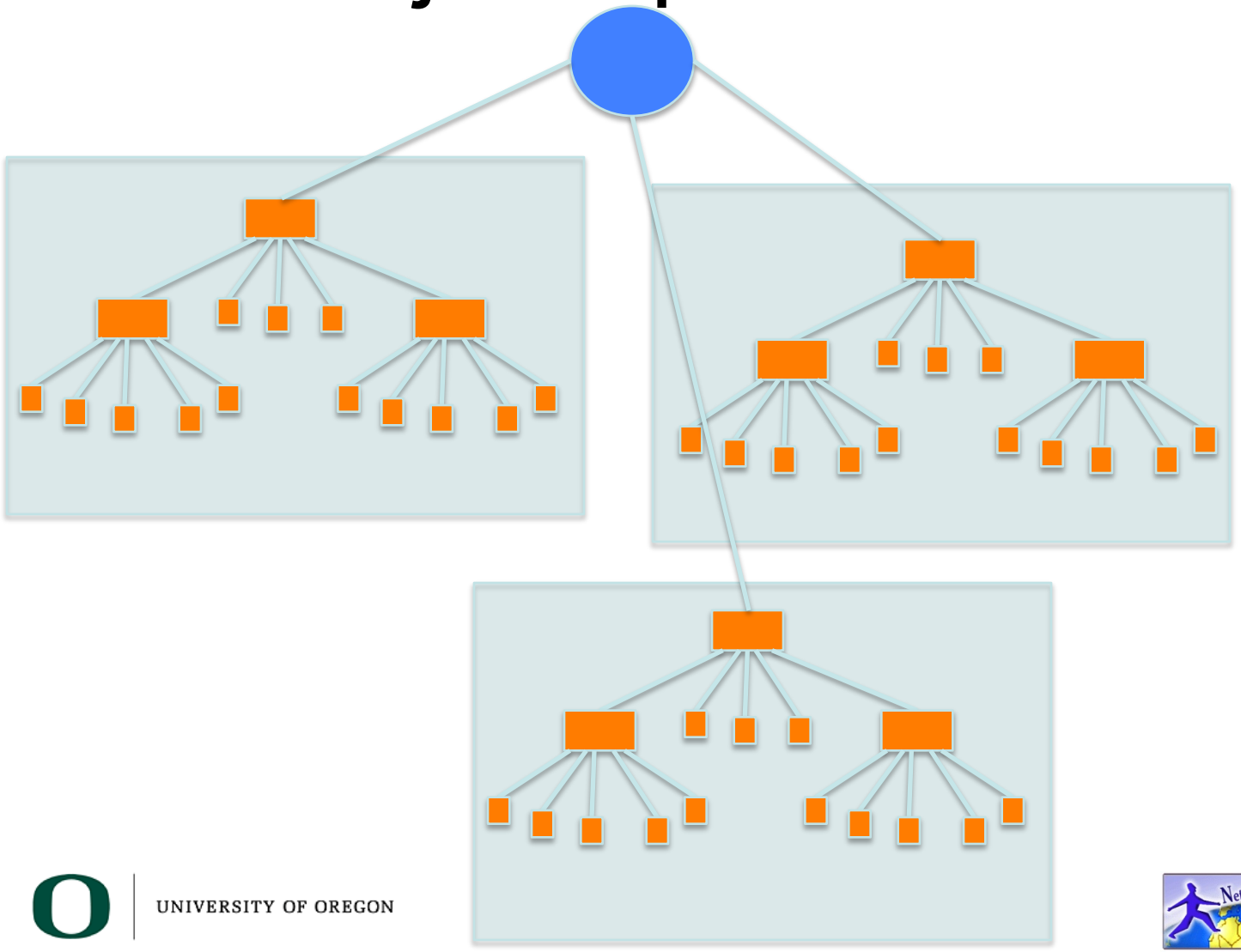


No encadene equipos

- Resista la tentación de hacer esto:

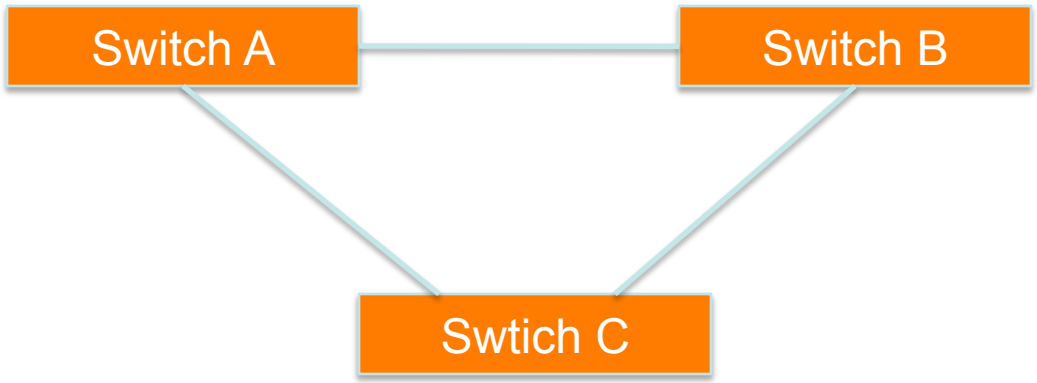


Conecte sus edificios jerárquicamente



Preguntas?

Bucle (loop) de capa 2

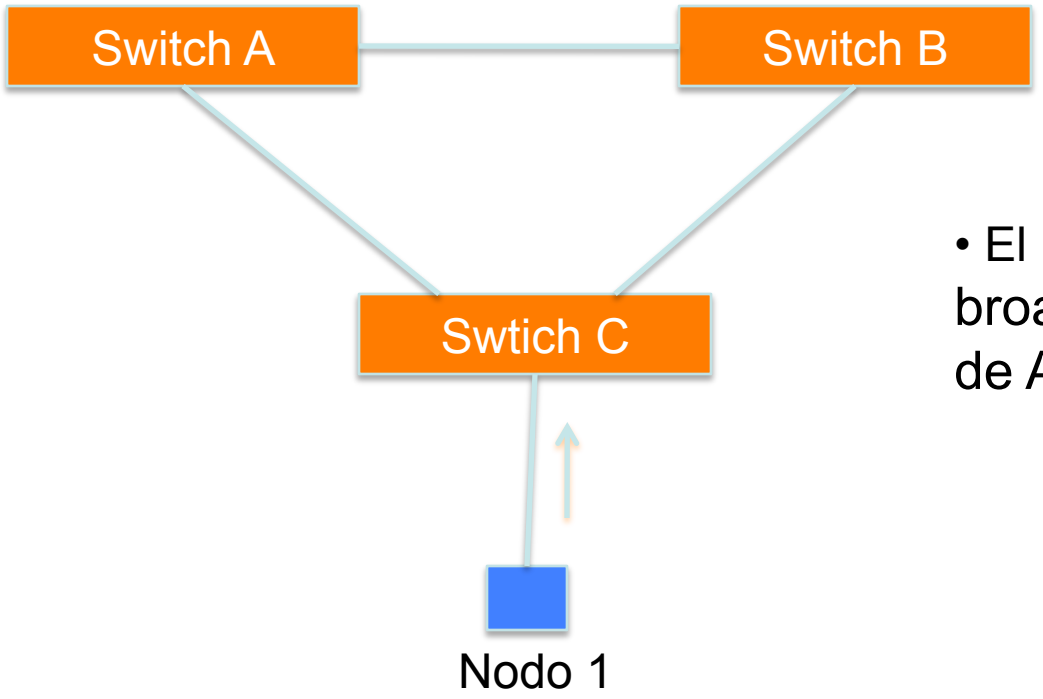


- Cuando hay más de un camino entre dos switches
- Cuáles son los posibles problemas?

Bucle de capa 2

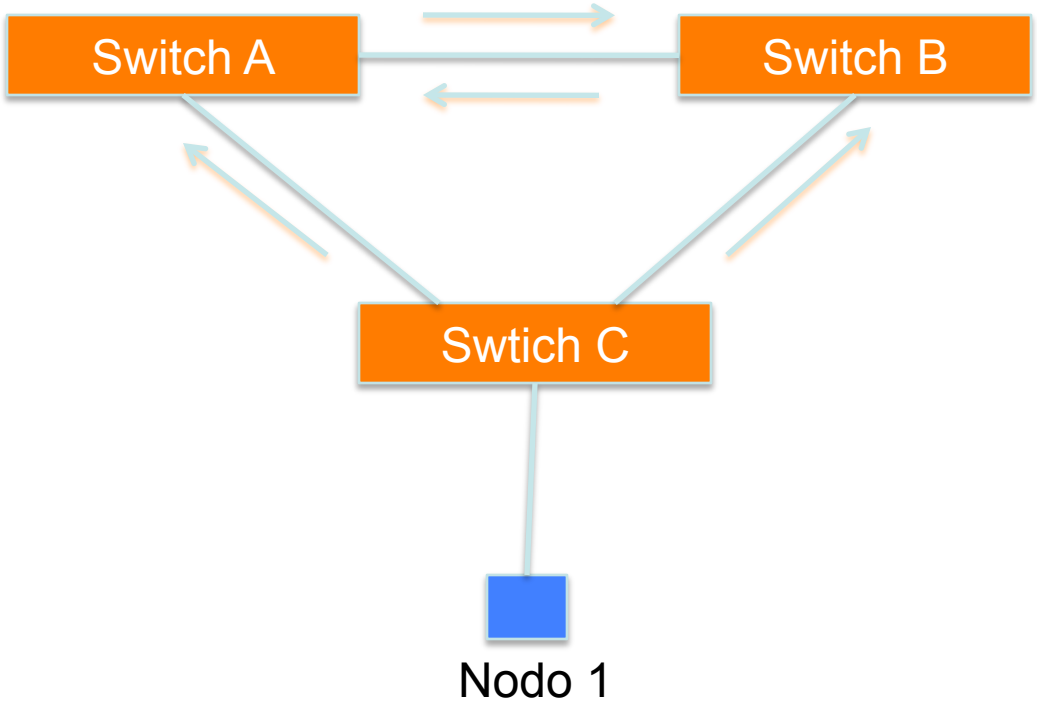
- Si hay más de un camino entre dos switches:
  - Las tablas de encaminamiento se hacen inestables
    - Las direcciones MAC de origen se ven venir intermitentemente desde puertos diferentes
  - Los switches se reenviarán los broadcasts entre sí
    - Todo el ancho de banda disponible será utilizado
    - Los procesadores de los switches no pueden con la carga

Bucle de capa 2



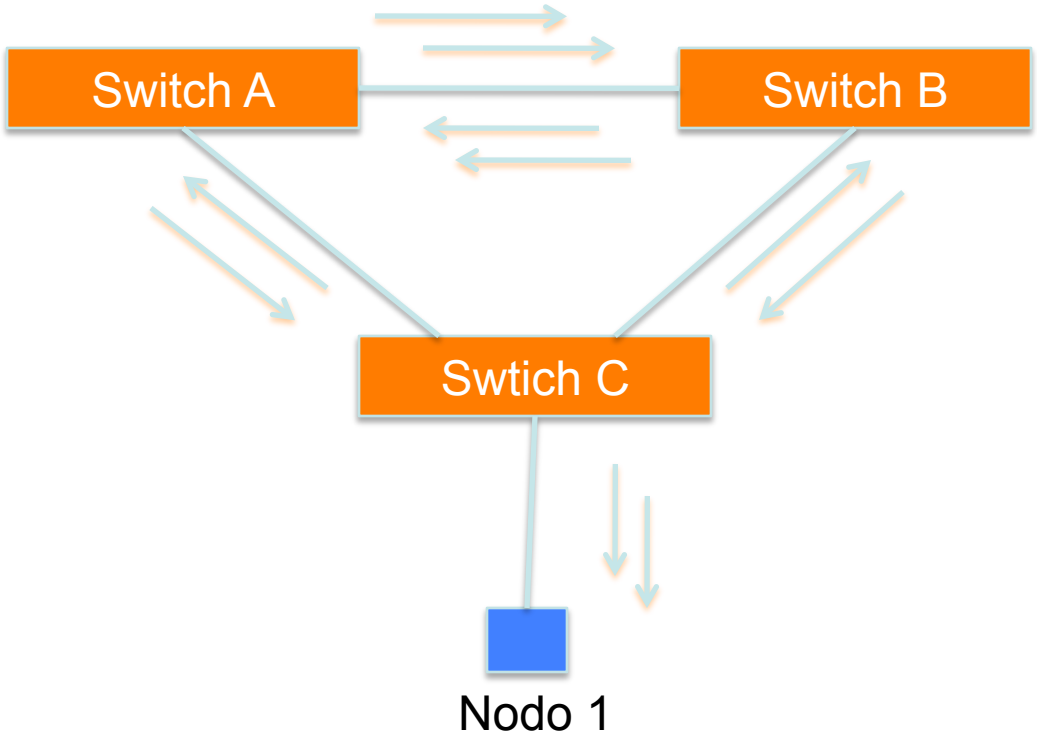
• El Nodo 1 envía una trama broadcast (ej. Una petición de ARP)

Bucle de capa 2



• Los switches A, B y C reenvían la trama del nodo 1 a través de todos los puertos

Bucle de capa 2



- Pero reciben sus propios broadcasts de nuevo, y pasan a reenviarlos otra vez!

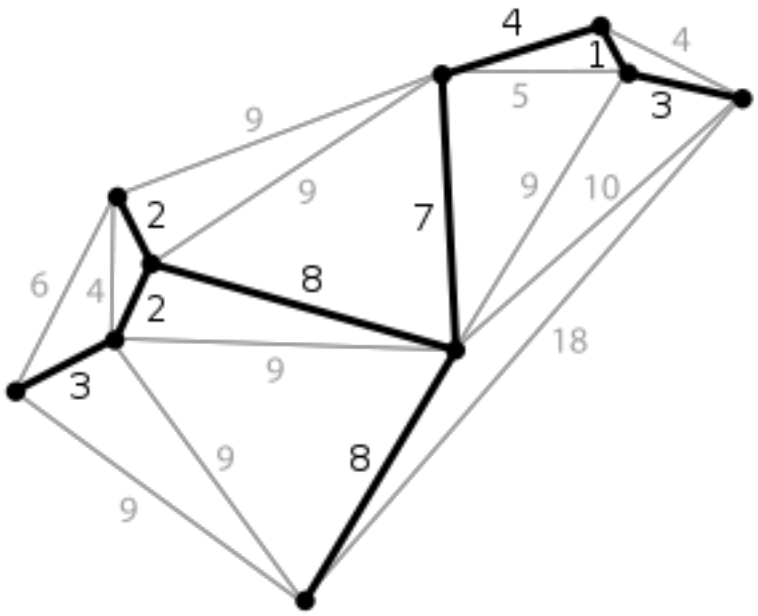
- Los broadcasts se amplifican, creando una **tormenta de broadcast**

Bucles buenos

- Se puede aprovechar de los bucles!
  - Los caminos redundantes mejoran la resistencia de la red cuando:
    - Un switch falla
    - Se rompe un enlace
- Cómo lograr redundancia sin crear bucles peligrosos entre switches?

Qué es un Spanning Tree

- “Dado un grafo conectado y sin dirección, un *spanning tree* de dicho grafo es un sub-grafo de tipo árbol que conecta todos los vértices”.
- Un solo grafo puede tener múltiples *spanning trees*.



Spanning Tree Protocol

- *El propósito del protocolo es hacer que los switches descubran de forma dinámica un subconjunto de la topología que esté libre de bucles (un árbol) y que aún tenga suficiente conectividad para que haya un camino entre cada switch, siempre que sea físicamente posible.*

Spanning Tree Protocol

- Varias versiones:
  - Traditional Spanning Tree (802.1d)
  - Rapid Spanning Tree o RSTP (802.1w)
  - Multiple Spanning Tree o MSTP (802.1s)

Traditional Spanning Tree (802.1d)

- Los switches intercambian mensajes que les permiten calcular el Spanning Tree
  - Estos mensajes se conocen como BPDUs (Bridge Protocol Data Units)
  - Dos tipos de BPDUs:
    - Configuración
    - Topology Change Notification (TCN)

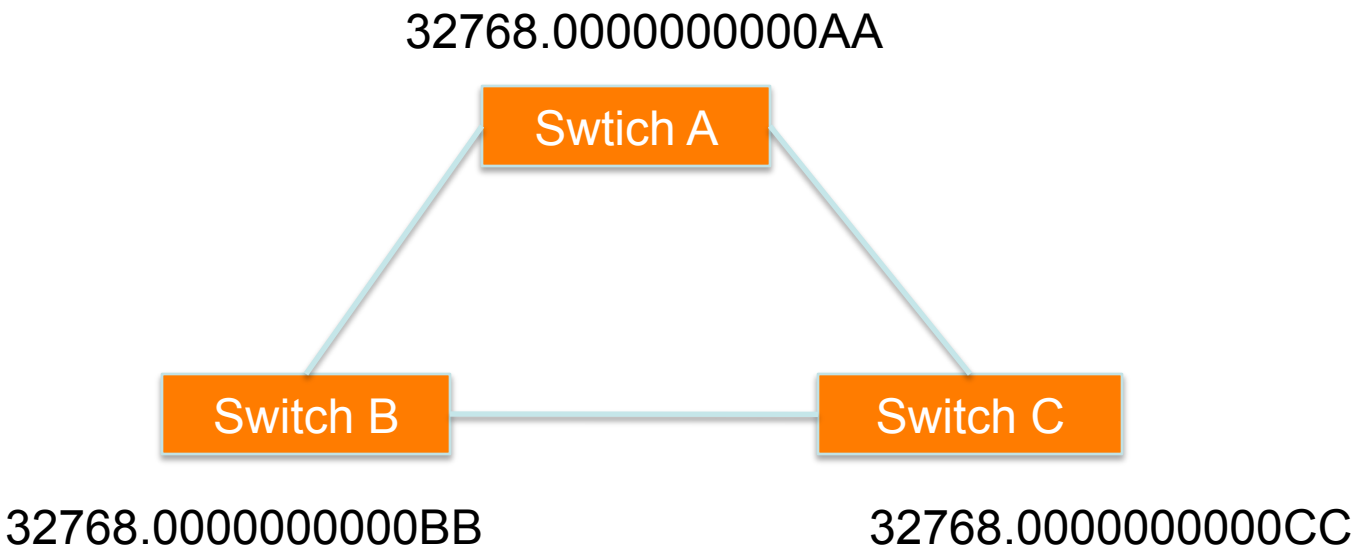
Traditional Spanning Tree (802.1d)

- Primer paso:
  - Decidir la ubicación del punto de referencia: el **switch raíz (root switch)**
  - El proceso de elección se basa en el ID del switch, que se compone de:
    - La prioridad del switch: Un valor de dos octetos que es configurable
    - La dirección MAC: Una dirección única, escrita en hardware, que no se puede cambiar.

Elección del switch raíz (802.1d)

- Cada switch comienza enviando BPDUs con un ID de switch raíz igual a su propio ID
  - *Yo soy el switch raíz!*
- Los BPDUs recibidos se analizan para ver si hay un ID de switch raíz que sea menor
  - De ser así, cada switch reemplaza el valor del ID del switch raíz anunciado con el valor menor
- Al cabo de un rato, todos los switches se ponen de acuerdo en quién será el switch raíz

Elección del switch raíz (802.1d)



- Todos los switches tienen la misma prioridad.
- Quién será elegido el switch raíz?

Selección del puerto raíz (802.1d)

- Ahora cada switch tiene que determinar dónde se encuentra en relación al switch raíz
  - Cada switch determina su ***Puerto Raíz***
  - La clave es encontrar el puerto con el menor ***Costo de camino a la raíz***
    - El costo acumulado de todos los enlaces que llevan al switch raíz

Selección del puerto raíz (802.1d)

- Cada enlace en cada switch tiene un **costo de camino (path cost)**
  - Inversamente proporcional a la capacidad del enlace
    - O sea, a mayor capacidad, menor costo

Capacidad de enlace	Costo de STP
10 Mbps	100
100 Mbps	19
1 Gbps	4
10 Gbps	2

Selección del puerto raíz (802.1d)

- ***El costo del camino a la raíz*** es la acumulación del costo de camino del puerto más los costos aprendidos de los switches vecinos.
  - Responde a la pregunta: *Cuánto cuesta alcanzar al switch raíz a través de este puerto?*

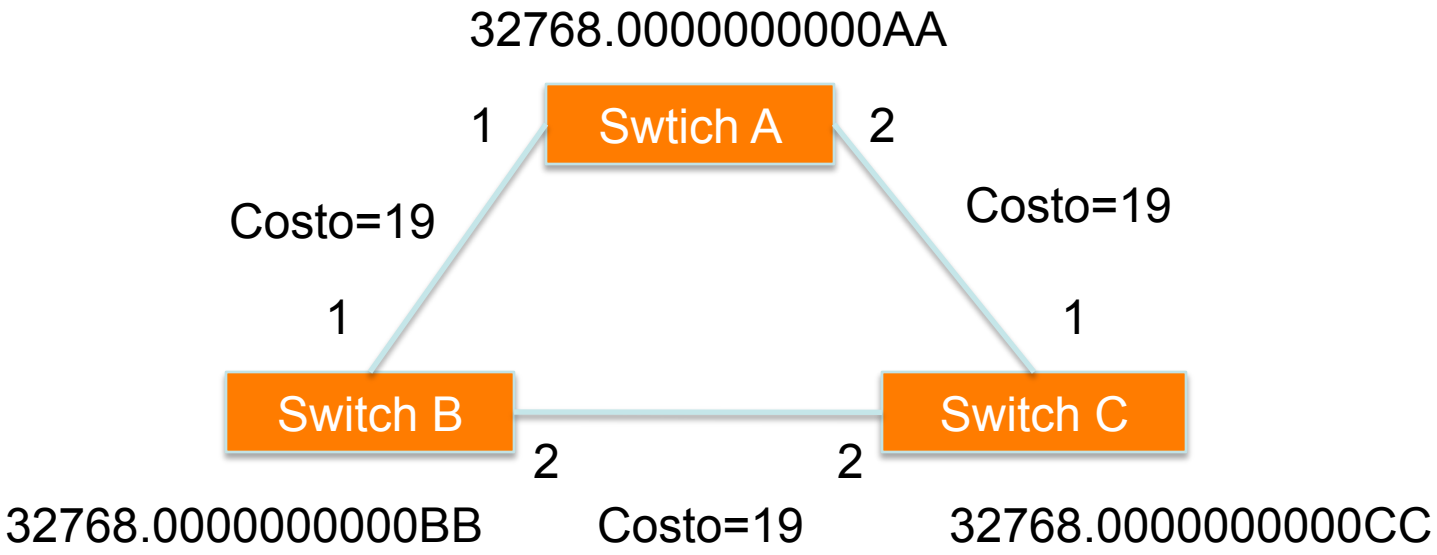
Selección del puerto raíz (802.1d)

1. El switch raíz envía BPDUs con un costo de camino a la raíz con valor 0
2. El switch vecino recibe el BPDU y agrega el costo del puerto al costo de camino a la raíz recibido
3. El switch vecino envía BPDUs con el nuevo valor acumulado
4. Cada vecino subsiguiene continúa la acumulación de la misma manera

Selección del puerto raíz (802.1d)

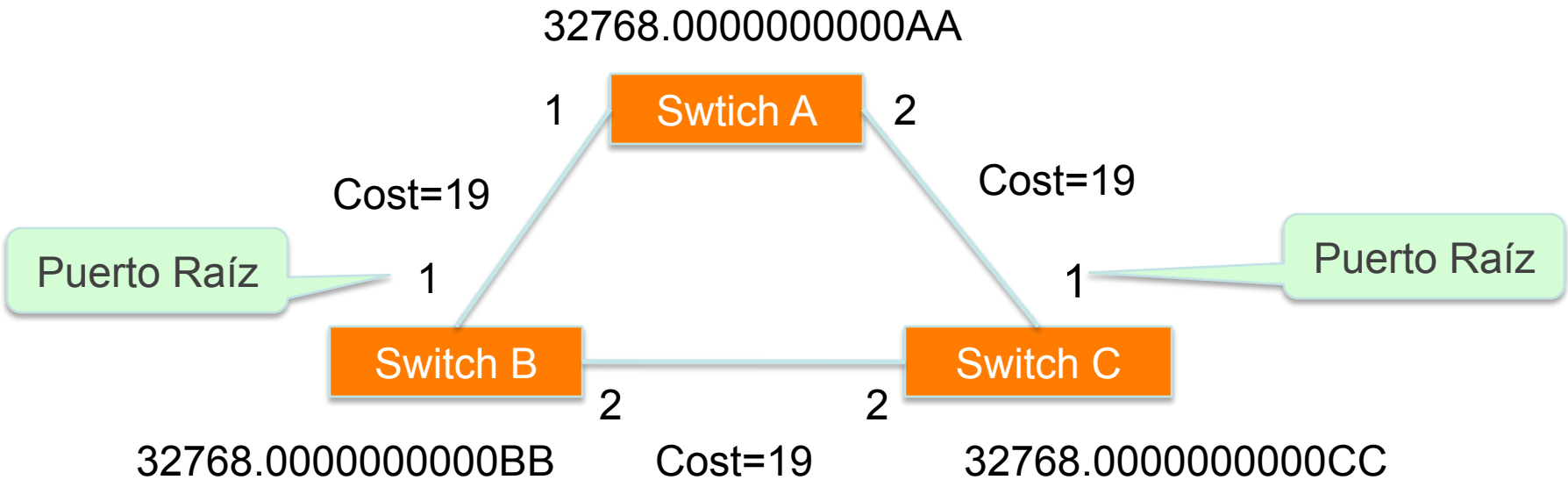
- En cada switch, el puerto donde se ha recibido el costo del camino a la raíz menor se designa como el ***Puerto Raíz***
  - Este es el puerto con el mejor camino al switch raíz

Selección del puerto raíz (802.1d)



- Cuál es el costo del camino a la raíz en cada puerto?
- Cuál es el puerto raíz en cada switch?

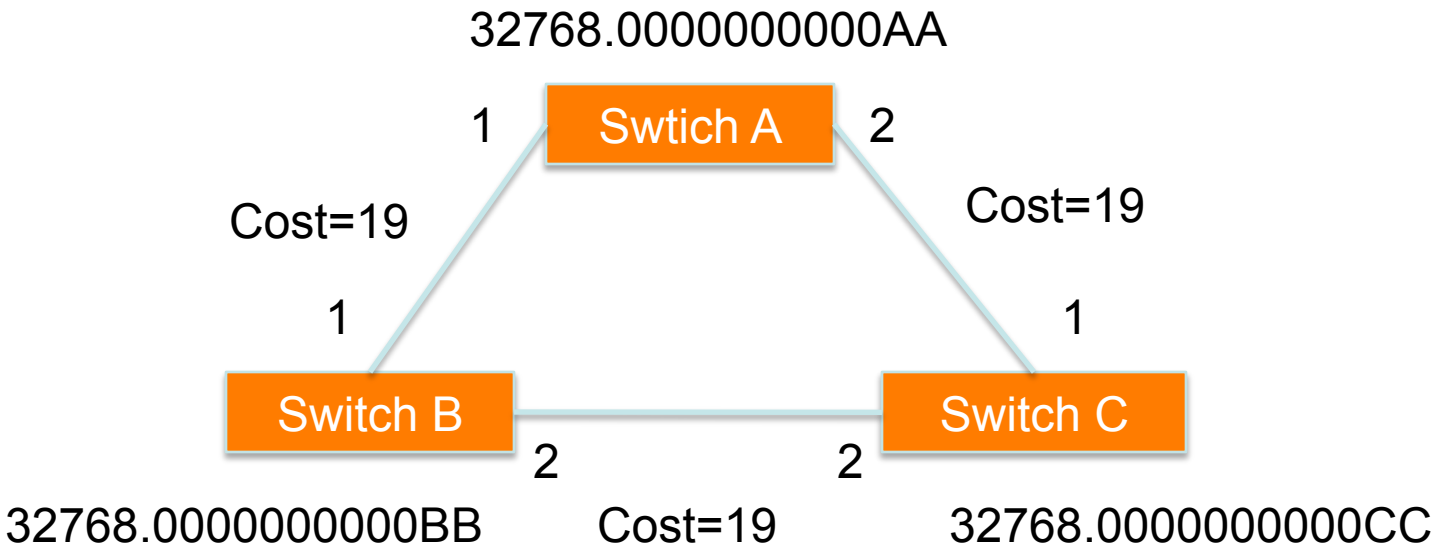
Selección del puerto raíz (802.1d)



Elección de puertos designados (802.1d)

- Bien, hemos seleccionado los puertos raíz, pero aún no hemos solucionado el problema
  - Los enlaces siguen activos!
- Cada segmento de red tiene que tener sólo un switch enviando tramas para ese segmento
- Cada switch tiene que identificar un ***Puerto designado*** por enlace
  - El enlace con el menor costo del camino a la raíz acumulado

Elección de puertos designados (802.1d)

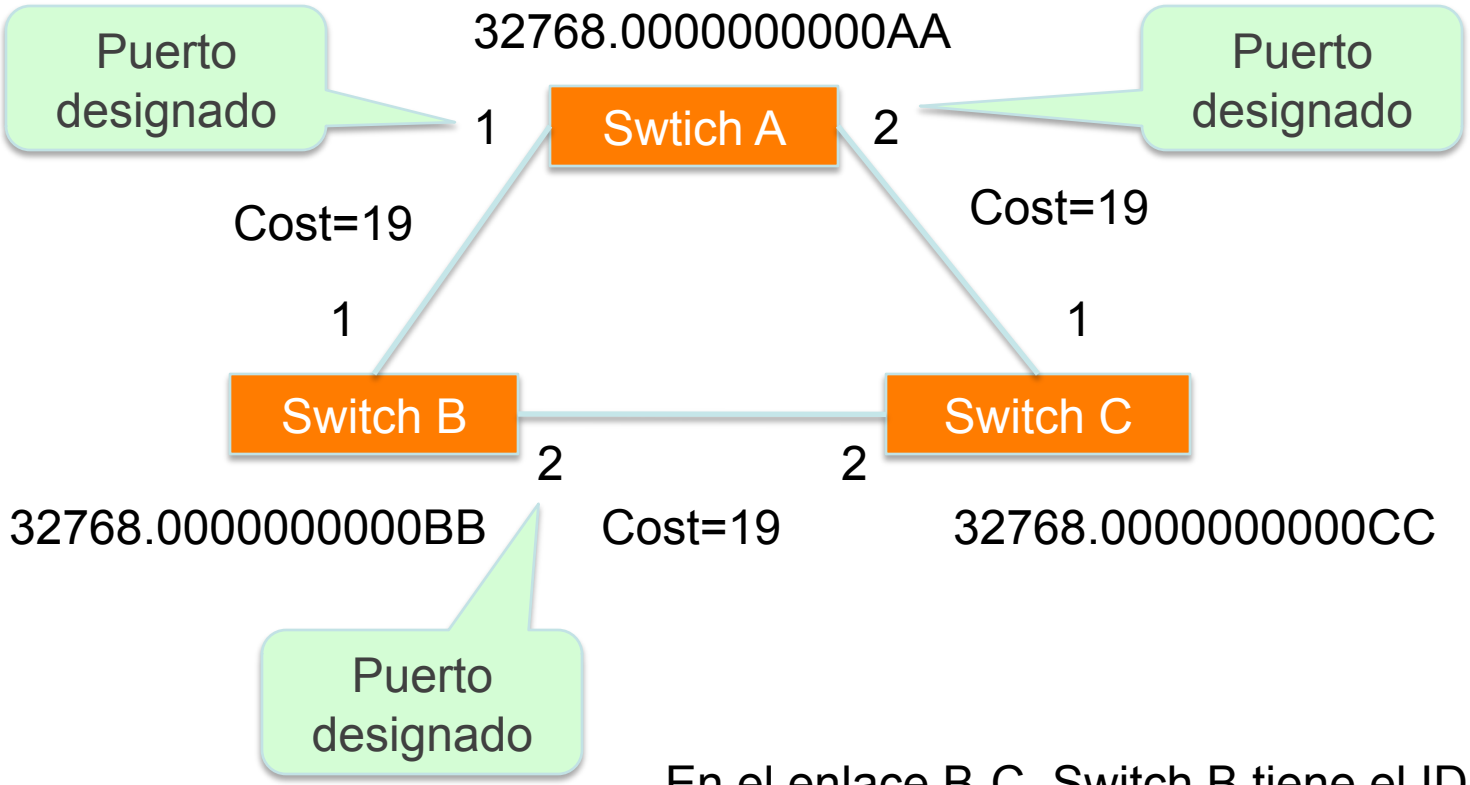


- Cuál puerto debe ser el puerto designado en cada segmento?

Elección de puertos designados (802.1d)

- Encontrar uno o más puertos en un segmento con costos de camino a la raíz es posible, lo cual resulta en un empate
- Todas las decisiones de STP están basadas en la siguiente secuencia de condiciones:
  - Menor ID de switch raíz
  - Menor costo del camino a la raíz
  - Menor ID de switch origen
  - Menor ID del puerto origen

Elección de puertos designados (802.1d)

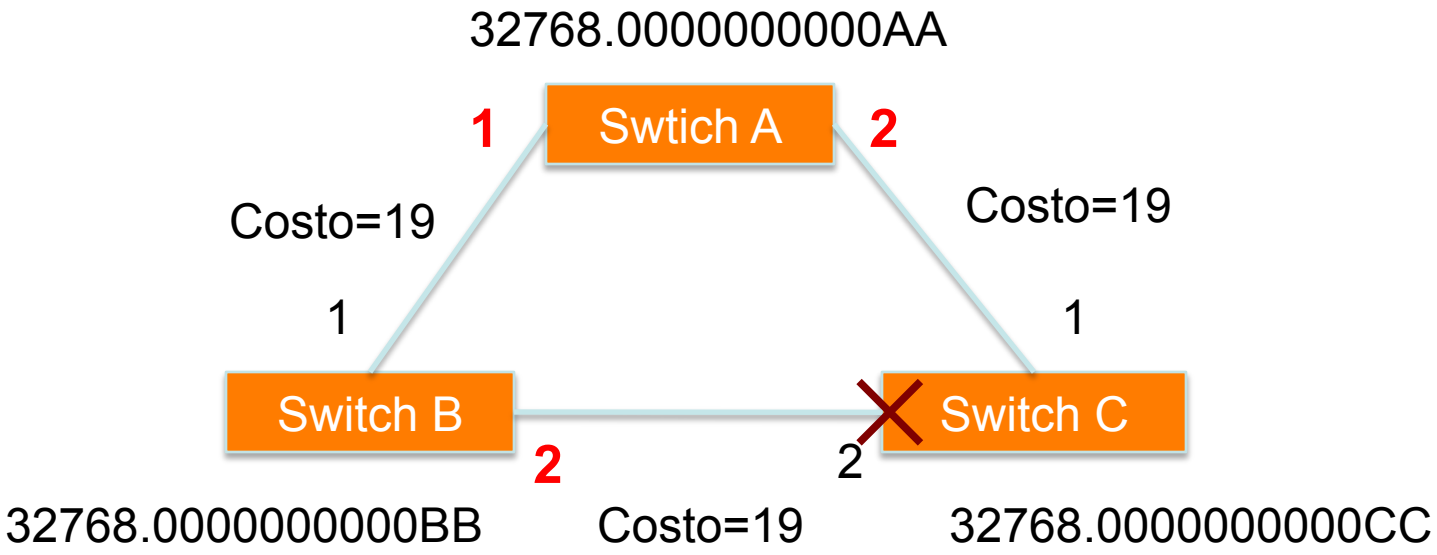


En el enlace B-C, Switch B tiene el ID menor, por lo que el puerto 2 en Switch B es el puerto designado

Bloqueo de puertos

- Cualquier puerto que no sea un puerto raíz o un puerto designado se pone en **estado bloqueado**
- Este paso efectivamente rompe el bucle y completa el Spanning Tree.

Puertos designados en cada segmento (802.1d)



- El Puerto 2 en Switch C se pone en ***Estado Bloqueado*** porque no es ni ***Puerto Raíz ni Puerto Designado***

Estados de Spanning Tree

- Desactivado (Disabled)
  - El puerto está apagado
- Bloqueado (Blocking)
  - Sin reenvío de tramas
  - Recibiendo BPDUs
- Escuchando (Listening)
  - Sin reenvío de tramas
  - Enviando y recibiendo BPDUs

Estados de Spanning Tree

- Aprendiendo (Learning)
  - Sin reenvío de tramas
  - Enviando y recibiendo BPDUs
  - Aprendiendo nuevas direcciones MAC
- Reenviando (Forwarding)
  - Reenviando tramas
  - Enviando y recibiendo BPDUs
  - Aprendiendo nuevas direcciones MAC

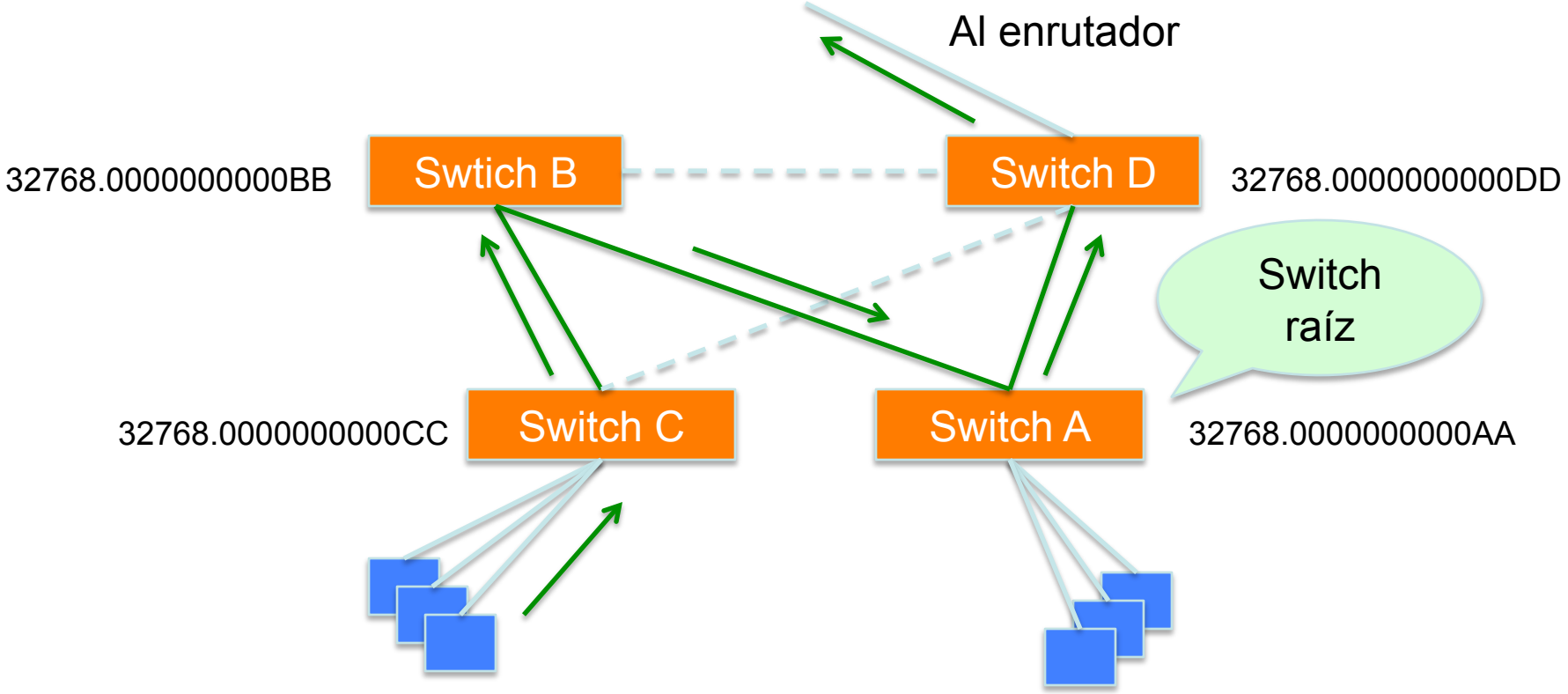
Cambios de Topología en STP

- Los switches recalculan si:
  - Se introduce un nuevo switch
    - Podría ser el nuevo raíz!
  - Un switch falla
  - Un enlace se cae

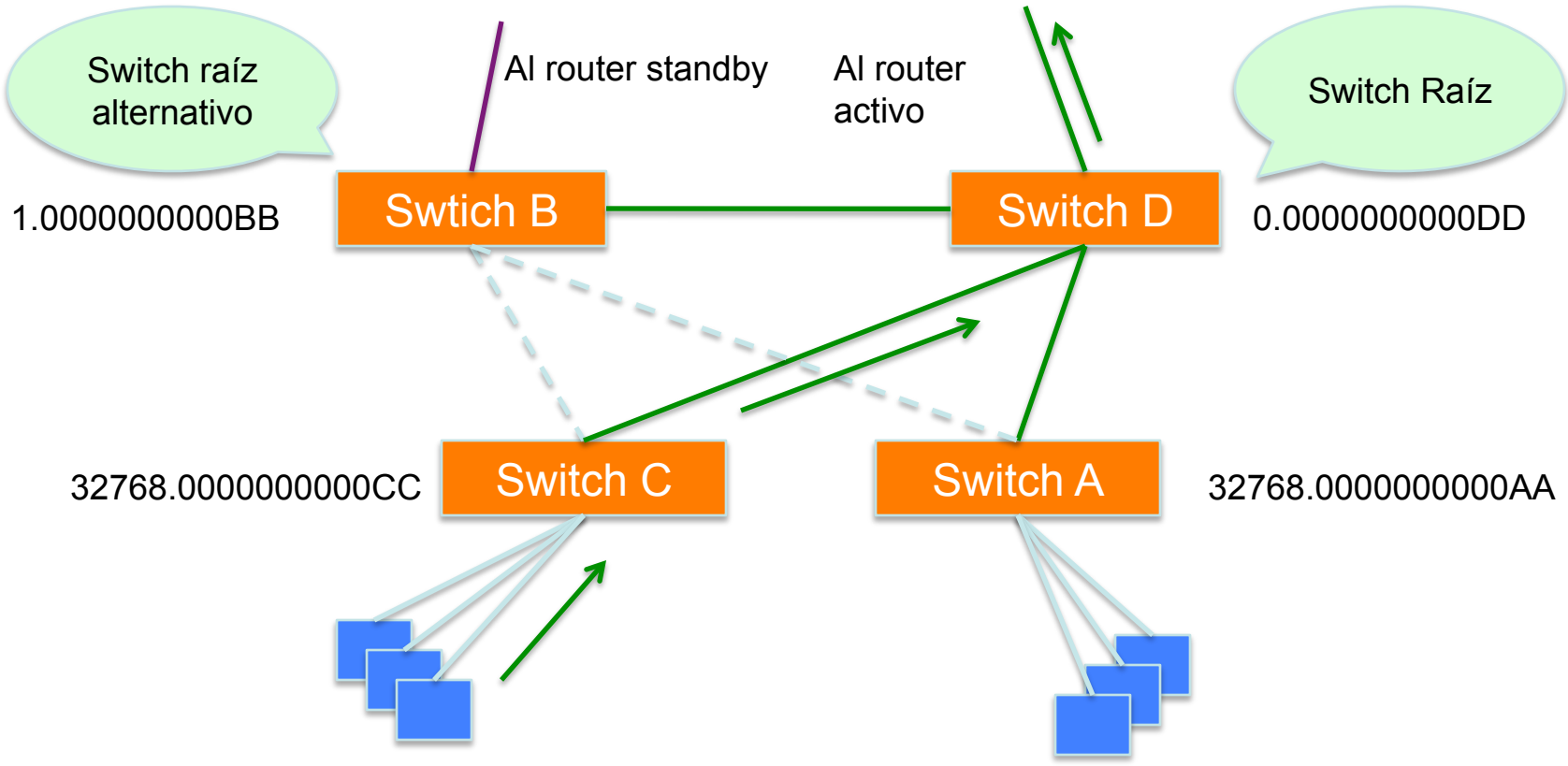
Ubicación del switch raíz

- Utilizar los parámetros por defecto puede resultar en una situación indeseada
  - El flujo de tráfico puede ser sub-óptimo
  - Un switch inestable o lento puede convertirse en el switch raíz
- Es necesario planificar la asignación de prioridades con cuidado

Mala ubicación del switch raíz



Buena ubicación del switch raíz



Protección de la topología STP

- Algunos fabricantes han introducido funcionalidades para proteger la topología:
  - Root Guard
  - BPDU Guard
  - Loop Guard
  - UDLD
  - Etc.

Pautas de diseño de STP

- Habilite el spanning tree aún si no tiene caminos redundantes
- Siempre planifique y asigne las prioridades
  - Haga la selección del switch raíz determinística
  - Incluya un switch raíz alternativo
- Si es posible, no acepte BPDUs en los puertos de los usuarios
  - Habilite BPDU Guard o similar donde esté disponible

Velocidad de Convergencia de 8021.d

- Cambiar del estado bloqueado al estado de reenvío se tarda por lo menos 2 x *Forward Delay* (~ 30 seg.)
  - Esto puede ser problemático al conectar máquinas de usuarios
- Algunos fabricantes han agregado mejoras como *PortFast*, el cual reduce el tiempo al mínimo en los puertos de usuarios
  - No use *PortFast* o similar en los enlaces entre switches
- Los cambios de topología también se tardan unos 30 segundos
  - Esto puede ser inadmisible en una red en producción

Rapid Spanning Tree (802.1w)

- La convergencia es mucho más rápida
  - La comunicación entre switches es más interactiva
- Los puertos de usuarios no participan
  - Estos van al estado de forwarding inmediatamente
  - Si se reciben BPDUs en un puerto de usuario, éste se convierte en un puerto inter-switch para evitar bucles

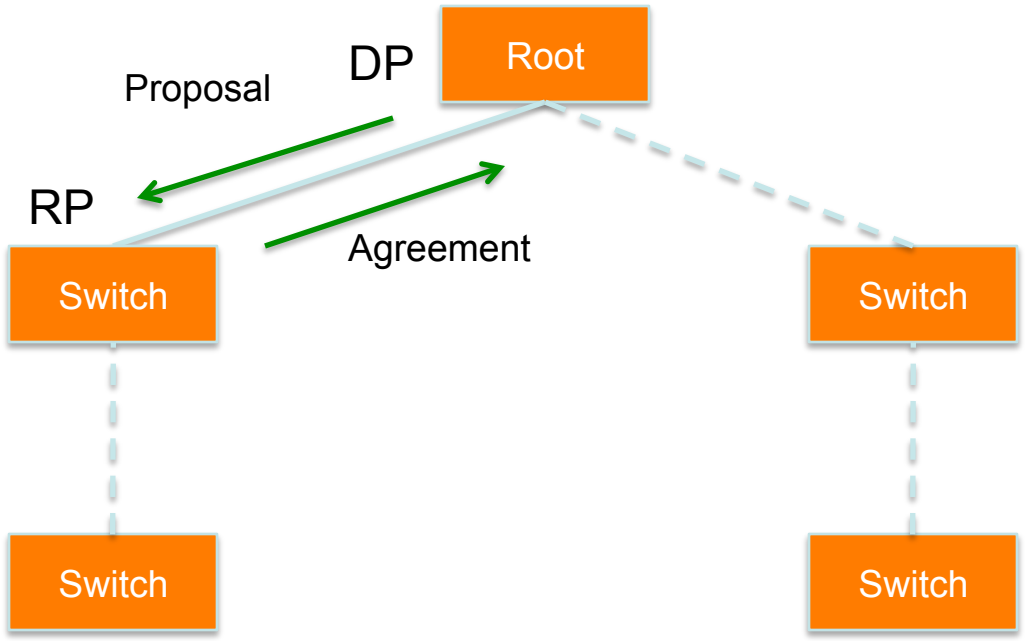
Rapid Spanning Tree (802.1w)

- Define estos roles de puerto:
  - Puerto Raíz (igual que en 802.1d)
  - Puerto Alternativo
    - Puerto con camino alternativo al switch raíz
  - Puerto Designado (igual que en 802.1d)
  - Puerto Backup
    - Camino backup/redundante a un segmento donde otro switch está conectado.

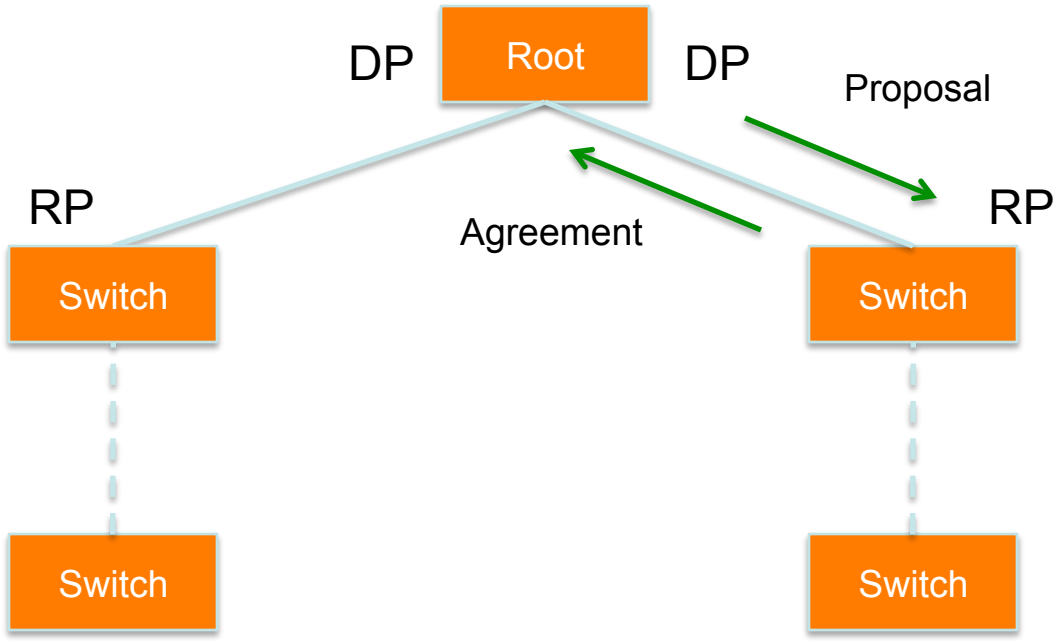
Rapid Spanning Tree (802.1w)

- El proceso de sincronización utiliza un método de *handshake*
  - Luego de elegirse el switch raíz, la topología se construye en cascada, donde cada switch propone ser el switch designado para cada enlace punto-a-punto
  - Mientras esto ocurre, todos los enlaces en los switches de niveles inferiores están bloqueados

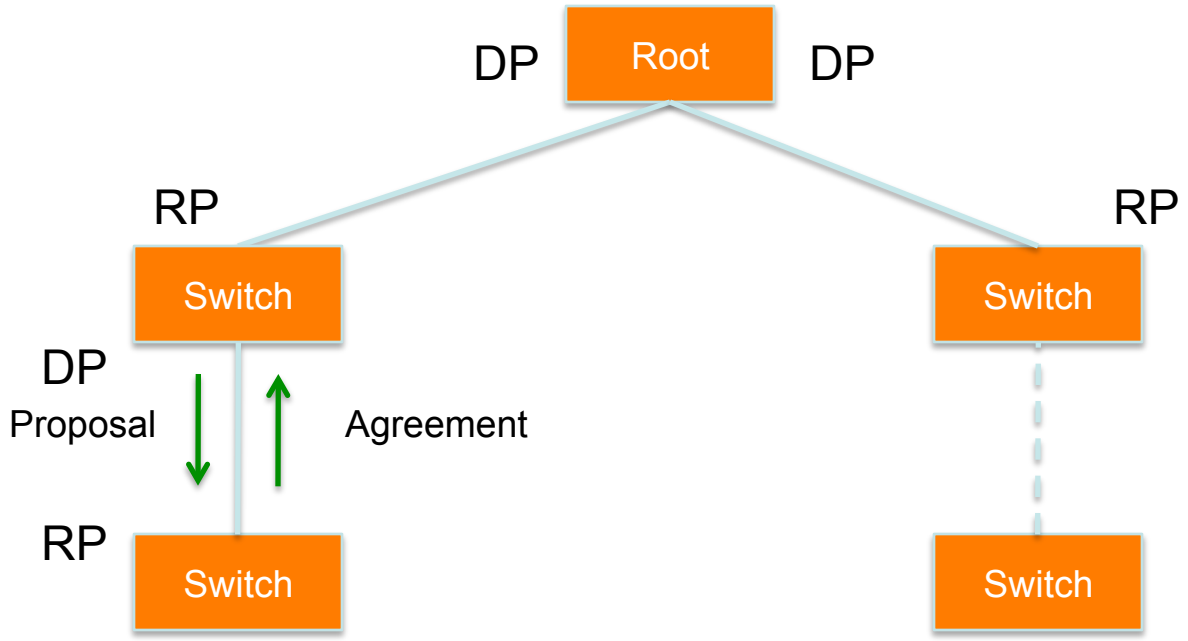
Rapid Spanning Tree (802.1w)



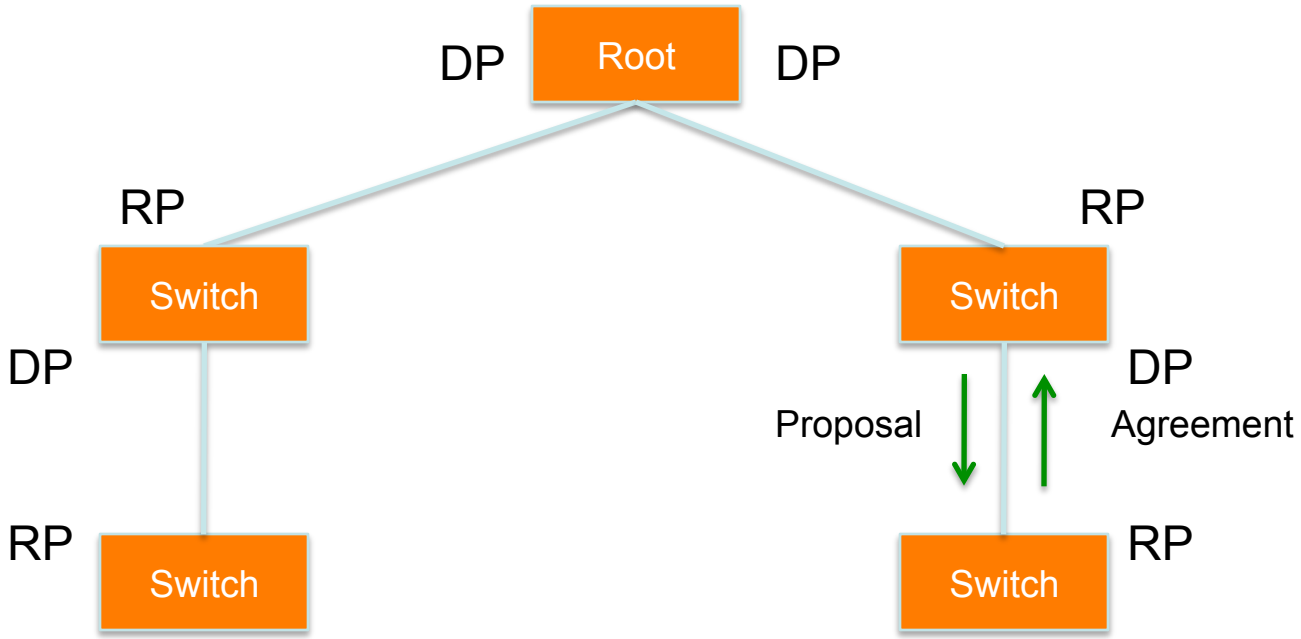
Rapid Spanning Tree (802.1w)



Rapid Spanning Tree (802.1w)



Rapid Spanning Tree (802.1w)



Rapid Spanning Tree (802.1w)

- Prefiera RSTP en lugar de STP si quiere convergencia más rápida
- Siempre defina cuáles son los puertos de los usuarios

Preguntas?

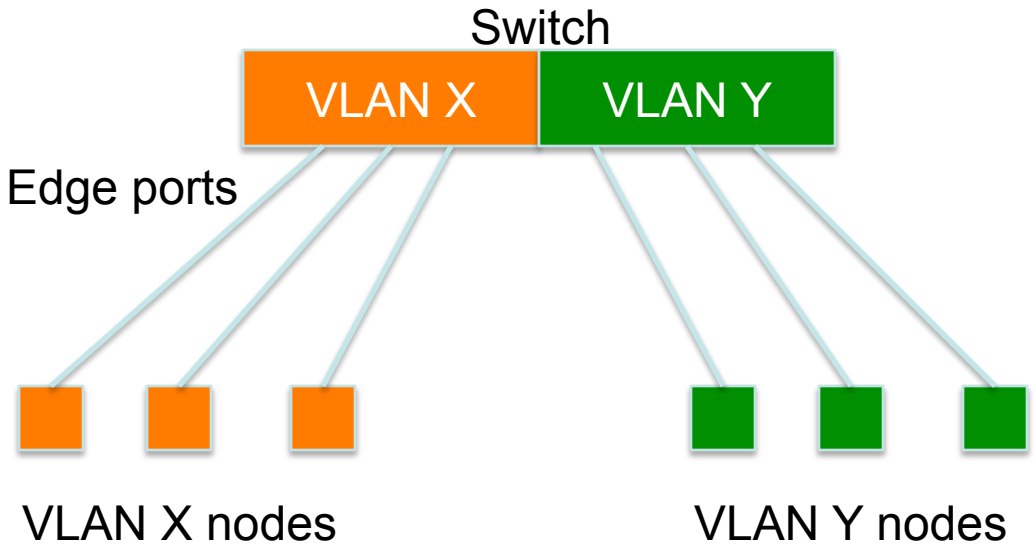
Virtual LANs (VLANs)

- Nos permiten separar los switches en varios switches virtuales
- Sólo los miembros de una VLAN pueden ver el tráfico de dicha VLAN
  - Tráfico entre VLANs debe pasar por un enrutador

VLANs locales

- 2 o más VLANs dentro de un mismo switch
- ***Los Puertos de usuario (Edge)***, donde las máquinas se conectan, se configuran como miembros de la VLAN
- El switch se comporta como varios switches separados, enviando tráfico solamente entre miembros de la misma VLAN

Local VLANs



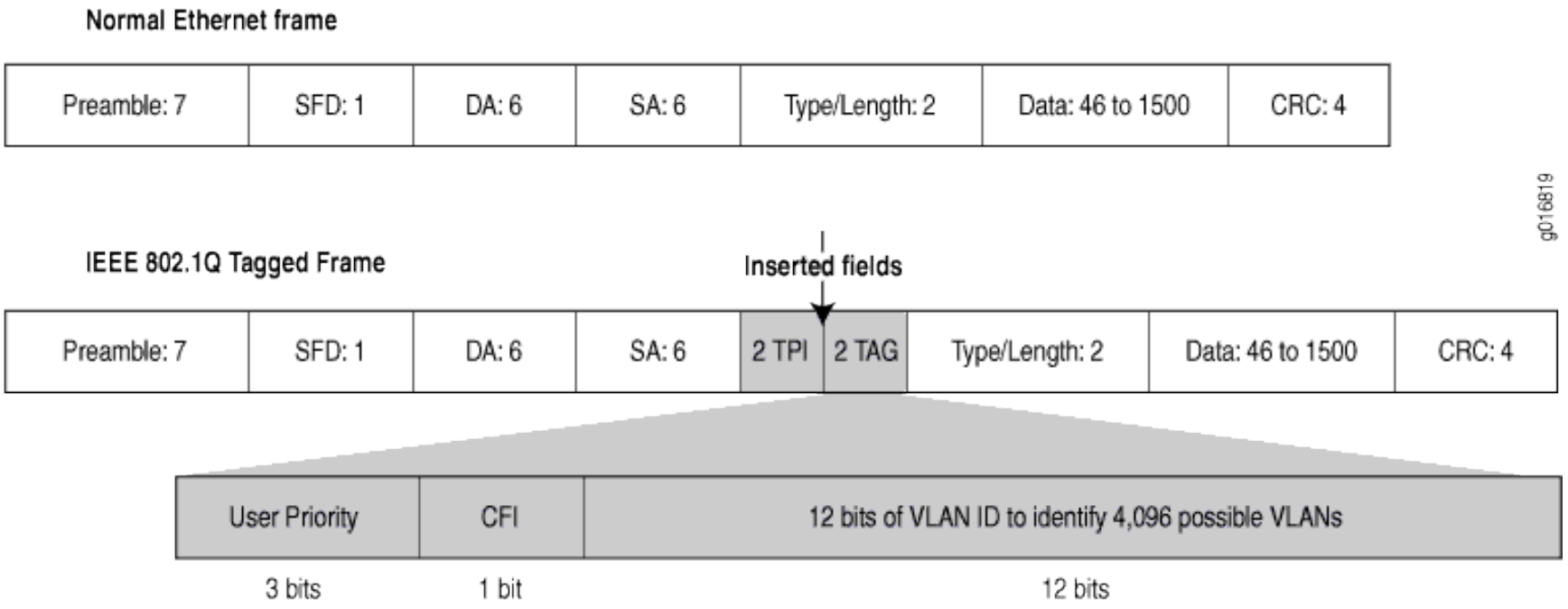
VLANs entre switches

- Dos o más switches pueden intercambiar tráfico de una o más VLANs
- Los enlaces inter-switch se configuran como ***troncales (trunks)***, transportando tramas de todas o una parte de las VLANs de un switch
- Cada trama lleva una **etiqueta** (***tag)*** que identifica la VLAN a la que pertenece

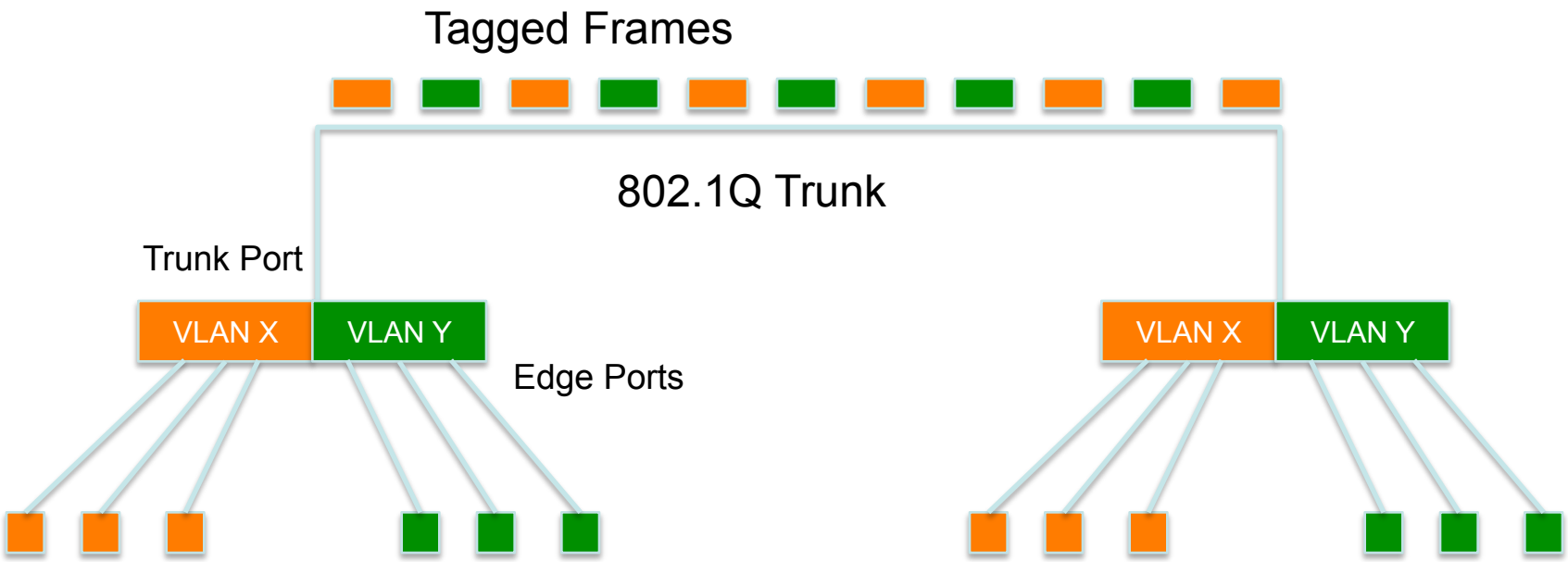
802.1Q

- El estándar de la IEEE que define cómo las tramas ethernet deberían ser etiquetadas ***tagged*** cuando viajan a través de troncales
- Esto implica que switches de **diferentes vendedores** son capaces de intercambiar tráfico entre VLANs

802.1Q tagged frame



VLANs entre switches



Esto se conoce como “VLAN Trunking”

Tagged vs. Untagged

- Los puertos de usuarios no se etiquetan, sólo se hacen “miembros” de una VLAN
- Sólo es necesario etiquetar tramas en puertos entre switches (*trunks*), cuando éstos transportan tráfico de múltiples VLANs
- Un *trunk* puede transportar tráfico de VLANs *tagged* y *untagged*
  - Siempre que los dos switches estén de acuerdo en cómo manejar éstas

Las VLANs aumentan la complejidad

- Ya no se puede simplemente “reemplazar” un switch
  - Ahora hay una configuración de VLANs que mantener
  - Los técnicos de campo necesitan más formación
- Hay que asegurarse de que todos los enlaces troncales están transportando las VLANs necesarias
  - Recordar cuando se esté agregando o quitando VLANs

Buenas razones para utilizar VLANs

- Hay que segmentar la red en varias subredes, pero no hay suficientes switches
- Separar los elementos de infraestructura como teléfonos IP, controles automáticos, etc.
- Separar el plano de control
  - Restringir quiénes puden acceder a la dirección de gestión del switch

Malas razones para usar VLANs

- Porque es posible, y le hace sentir “cool” ☺
- Porque le darán seguridad absoluta para sus usuarios (o así parece)
- Porque le permiten extender la red IP hasta otros edificios remotos
  - De hecho esto es muy común, pero es muy mala idea

No haga un “VLAN spaghetti”

- Extender una VLAN a través de múltiples edificios, o todo el campus
- Mala idea porque:
  - El tráfico broadcast viaja a través de todas las troncales, de un extremo al otro de la red
  - Una tormenta de broadcast se propagará a través de toda la extensión de la VLAN, y afectará las otras VLANS!
  - Una pesadilla para el mantenimiento y la resolución de problemas

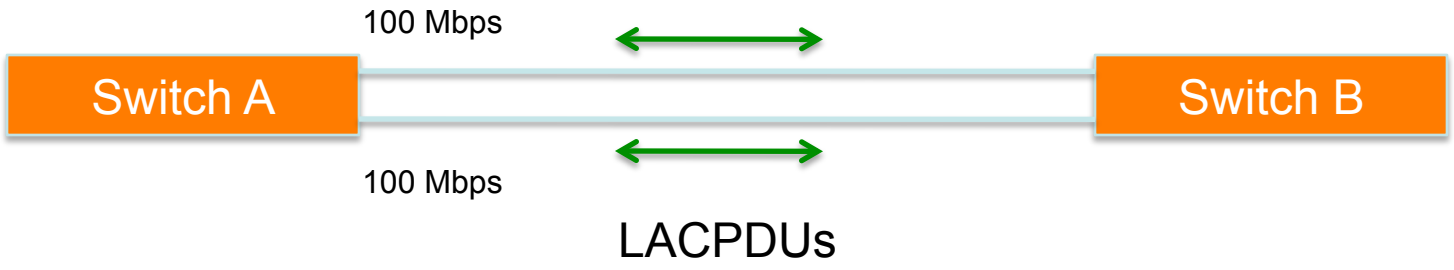
Agregación de Enlaces

- Conocido como *port bundling, link bundling*
- Se pueden usar varios enlaces en paralelo como si fueran un enlace único virtual
  - Para mayor capacidad del canal
  - Para redundancia (tolerancia a fallos)
- LACP (Link Aggregation Control Protocol) es un método estándar para negociar estos enlaces agregados entre switches

Operación de LACP

- Dos switches conectados via múltiples enlaces enviarán paquetes LACPDU, identificándose a sí mismos y a los puertos que los enlazan
- Entonces construirán los enlaces agregados y empezarán a pasar tráfico por ellos.
- Los puertos se pueden configurar como pasivos o activos

Operación de LACP

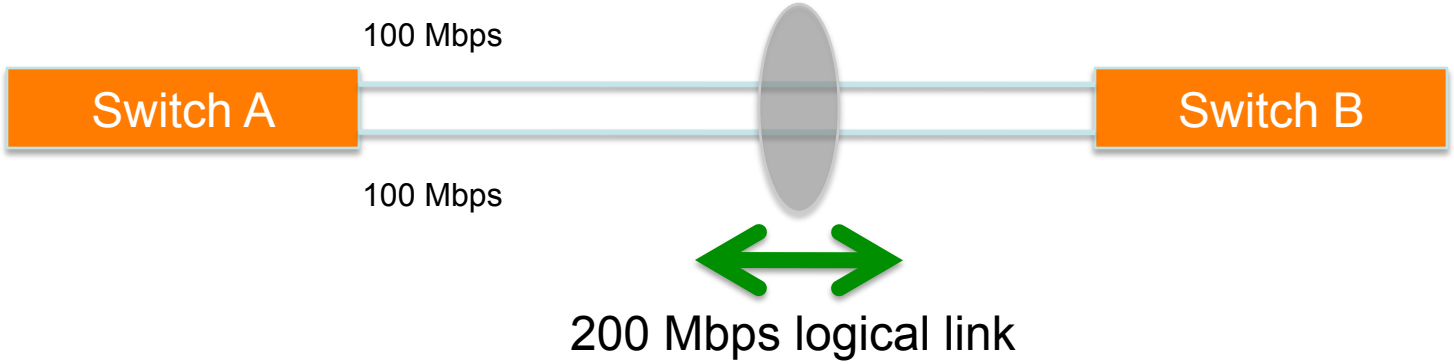


• Los switches A y B se conectan entre sí mediante dos pares de puertos Fast Ethernet

• LACP se habilita y los puertos se activan

• Los switches empiezan a enviar LACPDUs y negocian cómo establecer en enlace virtual

Operación de LACP



- El resultado es un enlace virtual agregado de 200 Mbps

- El enlace es también tolerante a fallos: Si uno de los enlaces miembro falla, LACP automáticamente quitará a ese enlace del grupo y seguirá enviando tráfico a través del enlace disponible

Distribución del tráfico en enlaces agregados

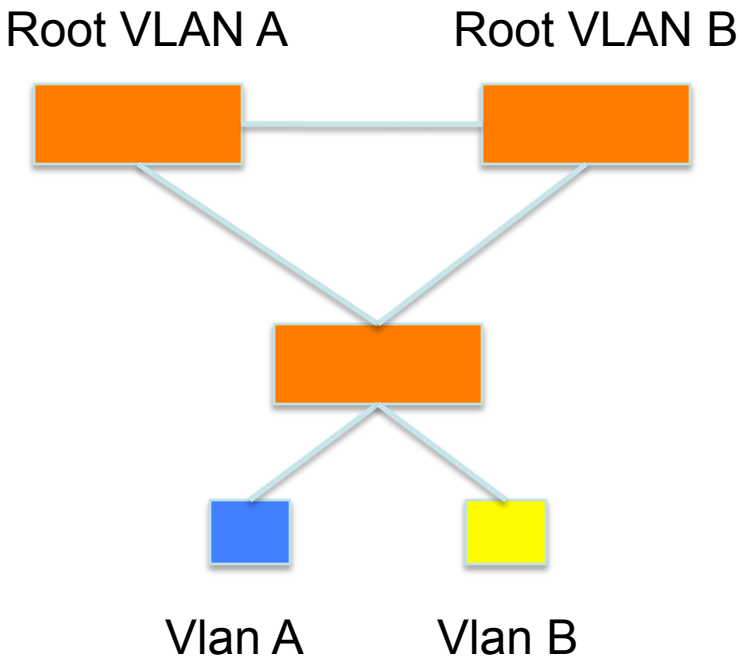
- Los enlaces agregados distribuyen las tramas gracias a un algoritmo, basado en:
  - Dirección MAC origen y/o destino
  - Dirección IP origen y/o destino
  - Números de puerto origen y/o destino
- Dependiendo de la naturaleza del tráfico, esto puede resultar en tráfico desbalanceado
- Siempre elija el método de balanceo de carga que provea la distrubución máxima

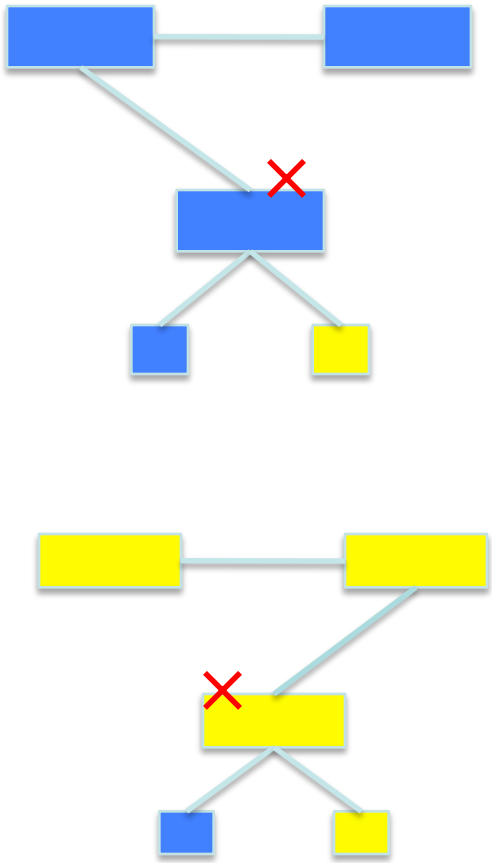
Preguntas?

Multiple Spanning Tree (802.1s)

- Permite crear “instancias” de Spanning Tree por cada grupo de VLANs
  - Las múltiples topologías permiten el balanceo de carga a través de diferentes enlaces
- Compatible con STP y RSTP

Multiple Spanning Tree (802.1s)





Multiple Spanning Tree (802.1s)

- Región MST
  - Los switches son miembros de una misma región si coinciden en sus parámetros:
    - Nombre de configuración MST
    - Número de revisión de la configuración MST
    - Mapeo de VLANs a instancias
  - Un resumen hash de estos atributos se envía dentro de las BPDUs para su rápido análisis en los switches
  - Una región es generalmente suficiente

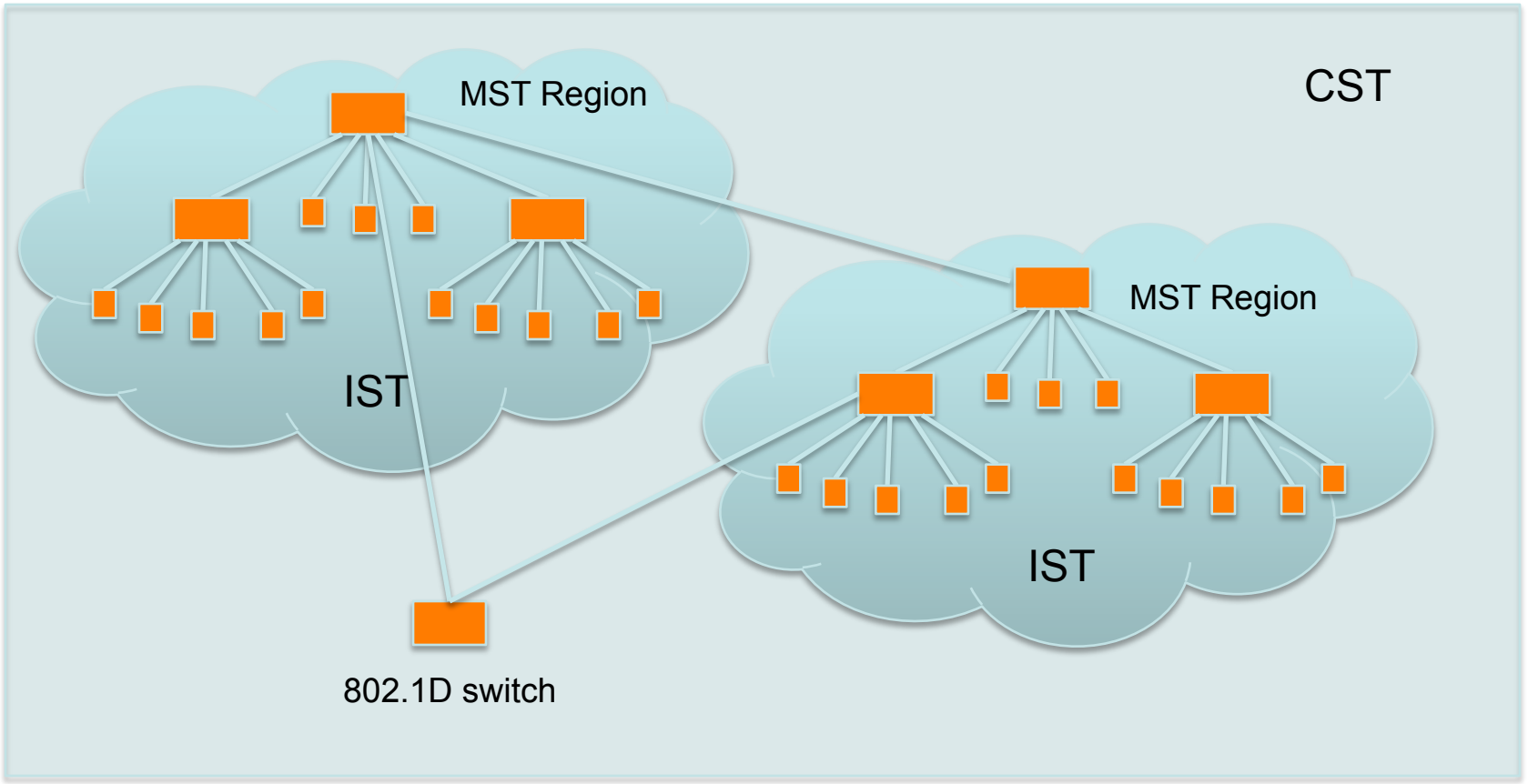
Multiple Spanning Tree (802.1s)

- CST = Common Spanning Tree
  - Para interoperar con otras versiones de Spanning Tree, MST necesita un spanning tree común que contenga todas las demás “islas”, incluyendo otras regiones MST

Multiple Spanning Tree (802.1s)

- IST = Internal Spanning Tree
  - Interno a la región
  - Presenta toda la región como un switch virtual único al CST externo

Multiple Spanning Tree (802.1s)



Multiple Spanning Tree (802.1s)

- Instancias MST
  - Groupos de VLANs se mapean a distintas instancias de MST
  - Estas instancias representarán cada topología alternativa, o caminos de reenvío alternativos
  - Se especifica un switch raíz y uno alternativo para cada instancia

Multiple Spanning Tree (802.1s)

- Pautas de diseño
  - Determinar los caminos de reenvío relevantes y distribuir las VLANs de manera equitativa entre las instancias correspondientes a cada uno de estos caminos
  - Designar los switches raíz y alternativo para cada instanacia
  - Asegurarse de que todos los switches concuerdan en sus parámetros
  - No asignar VLANs a la instancia 0, ya que ésta es utilizada por el IST

Elección de Switches

- Funcionalidades mínimas:
  - Conformidad con los estándares
  - Gestión cifrada (SSH/HTTPS)
  - VLAN trunking
  - Spanning Tree (por lo menos RSTP)
  - SNMP
    - Por lo menos versión 2 (v3 tiene mejor seguridad)
    - Traps

Elección de Switches

- Otras funcionalidades recomendadas:
  - DHCP Snooping
    - Evitar que sus usuarios activen un servidor DHCP ilegítimo
      - Ocurre mucho con los enrutadores wireless de bajo coste (Netgear, Linksys, etc) enchufados al revés
    - Los puertos que suben hasta el servidor DHCP legítimo se designan como “trusted”. Si hay DHCPOFFERs originadas desde puertos no confiados, son descartadas.

Elección de Switches

- Otras funcionalidades recomendadas:
  - Inspección de ARP dinámica
    - Un nodo malicioso puede realizar un ataque “man -in-the-middle” al enviar respuestas ARP ilegítimas
    - Los switches pueden mirar dentro de los paquetes ARP y descartar los que no sean legítimos.

Selección de Switches

- Otras funcionalidades recomendadas:
  - IGMP Snooping:
    - Los switches por defecto reenvían las tramas multicast a través de todos sus puertos
    - Al "husmear" el tráfico IGMP, el switch puede aprender cuáles máquinas son miembros de un grupo multicast, y enviar las tramas a través de los puertos necesarios solamente
    - Muy importante cuando los usuarios utilizan Norton Ghost, por ejemplo.

Gestión de Red

- Habilite los SNMP traps y/o Syslog
  - Reunir y procesar en un servidor central:
    - Cambios de Spanning Tree
    - Discordancias de Duplex
    - Problemas de cableado
- Monitorizar las configuraciones
  - Usar RANCID para reportar todos los cambios que ocurran en la configuración del switch

Gestión de Red

- Reuna y guarde las tablas de reenvío usando SNMP periódicamente
  - Le permite encontrar las direcciones MAC en su red de forma rápida
  - Puede usar archivos de texto simple y buscar con grep, o usar una herramienta con interfaz web y una base de datos
- Active LLDP (o CDP o similar)
  - Le muestra cómo los switches están interconectados entres sí, y a otros dispositivos

Documentación

- Documente la ubicación de sus switches
  - Nombre el switch basado en su ubicación
    - E.g. edificio1-sw1
  - Mantenga un récord de la ubicación física
    - Nivel, número de clóset, etc.
- Documente las conexiones a las tomas de red
  - Número de salón, número de toma, nombre del servidor, printer, etc.

Preguntas?

- Gracias.