Bachelor’s Thesis

Bachelor Degree in Information Technology

NINFOS10

2016

Instructor: Väänänen Ossi

Madhu Sudan Guragain

EVALUATING THE FUTURE OF THE SPANNING TREE PROTOCOL

TURKU AMK

TURKU UNIVERSITY OF APPLIED SCIENCES

BACHELOR’S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Bachelor Degree in Information Technology

2016 | 33

Madhu Sudan Guragain

EVALUATING THE FUTURE OF THE SPANNING TREE PROTOCOL

With the advancement in the technology in the present world, the limitations of the Spanning Tree Protocol in networking are creating increasing problems in Network Virtualization. The demand for smooth networking in big mess topology, its scalability and reliability are not met by the Spanning Tree Protocols (STPs). Transparent Interconnections of Lots of Links (TRILL) and Shortest Path Bridging (SPB) nowadays are becoming reliable and practical protocols in solving this problem in the networking industry.

To deploy either TRILL or SPB has been emerging as a challenging and hot debate for many large vendors such as Arista, Avaya, Brocade, Cisco, Extreme, and HP. This thesis explains the limitations on STPs regarding Virtualized Data Centers and Cloud Computing, and elaborates on the use, operation, and the differences of both TRILL or SPB.

KEYWORDS:

Spanning Tree Protocol, Spanning Tree Algorithm, Networking, TRILL, SPB, Redundancy.

CONTENTS

LIST OF ABBREVIATIONS......5

1. INTRODUCTION......6

2. SPANNING TREE PROTOCOL......7

2.1. STP Operation...... 7

2.1.1. Spanning Tree Algorithm: Port Roles...... 8

2.1.2. Spanning Tree Algorithm: Root Bridge...... 9

2.1.3. Spanning Tree Algorithm: Path Cost......10

2.2. BPDU Frame......10

2.3. Extended System ID......11

3. TYPES OF SPANNING TREE PROTOCOL......13

3.1. STP ...... **Error! Bookmark not defined.**

3.2. PVST+......13

3.3. 8021d-2004 3.3 13

3.4. RSTP (IEEE802.1w)......13

3.5. RPVST+ 13

3.6. MSTP 13

4. LIMITATIONS AND CHALLENGES OF SPANNING TREE......15

4.1. Spanning Tree Failure 15

4.1.1. Spanning Tree Convergence 15

4.1.2. Duplex Mismatch 15

4.1.3. Unidirectional Link......16

4.1.4. Packet Corruption 17

4.1.5. Resource Errors......17

4.1.6. PortFast Configuration Error 17

4.1.7. Awkward STP Parameter Tune and Diameter Issues 18

4.1.8. Software Errors......19

4.2. Need of Simplifying Network Virtualization......19

5. REPLACEMENT OF SPANNING TREE PROTOCOLS......21

5.1. IETF TRILL – Transparent Interconnection of Lots of Links 22

5.2. Shortest Path Bridging (SPB) 25

5.2.1. Shortest Path Bridging VID- SPBV......25

5.2.2. Shortest path bridging MAC- SPBM......26

5.3. Comparison of TRILL and SPB..28
6. CONCLUSION..33
REFERENCES..35

FIGURES

Figure 1. STP Operation: how STP compensate a network failure. 8
Figure 2. Port roles 9
Figure 3. Path to the root bridge...10
Figure 4. BPDU propagation and process...11
Figure 5. Extended system ID..12
Figure 6. Spanning tree protocol characteristics...14
Figure 7. Duplex mismatch...16
Figure 8. Unidirectional link..17
Figure 9. PortFast configuration error...18
Figure 10. Network virtualization in data center..20
Figure 11. Limitations on Spanning Tree Protocol...21
Figure 12. Encapsulation and Decapsulation in Trill...23
Figure 13. Ethernet and Trill Headers 24
Figure 14. Mapping of C-VLAN 20 over I-SID 100 and S-Tag bundled Interface In SPBM........28
Figure 15. TRILL and SPB lookup/forwarding 31
Figure 16. Shortest Path Trees 32

LIST OF ABBREVIATIONS

BID	Bridge ID
BPDU	Bridge Protocol Data Unit
CPU	Central Processing Unit
CDP	Cisco Discovery Protocol
CIST	Common and Internal Spanning Tree
ECT	Equal Cost Tree
ID	Identity
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IOS	Internetwork Operating System
IS-IS	Intermediate System to Intermediate System
MAC	Media Access Control
MLAG	Multi System Link Aggregation
MSTP	Multiple Spanning Tree Protocol
OSI	Open System Interconnection
OS	Operating System
PVST	Per-VLAN Spanning Tree
RPFC	Reverse Path Forwarding Check
RPVST	Rapid PVST
RSTP	Rapid Spanning Tree Protocol
SPB	Shortest Path Bridging
SPBM	Shortest Path Bridging MAC
SPBV	Shortest Path Bridging VLAN Identifier
SPT	Shortest Path Tree
SPVID	Shortest Path Vlan Identifier
STA	Spanning Tree Algorithm
STP	Spanning Tree Protocol
TRILL	Transparent Interconnection of Lots of Links
TTL	Time to Leave
UDLD	Unidirectional Link Detection
VID	VLAN Identifier
VLAN	Virtual LAN
WG	Working Group

1. INTRODUCTION

The Spanning Tree Protocol (STP) functions based on the spanning tree algorithm and its primary objective is to cut the loops and provide a redundant network topology. With the advancement of technology and requirement of thousands of devices being connected in the same topology, the traditional STP fails to deliver its services. A failure in the STA creates a bug in the network resulting in a bridging loop.

Thus, various alternatives for the Spanning Tree Protocols are being searched in the networking industry at the present. In the scenario that large private companies are having their own cloud computing environments, networking virtualization has become prominent and a key in creating an overall win-win situation for the companies. The question is now whether these traditional Spanning Tree protocols and their different variants like Per-VLAN Spanning Protocol (PVST), Rapid Spanning Tree Protocol (RSTP) and Multiple Spanning Tree Protocol (MSTP) are able to deliver the services required for the future.

Transparent Interconnection of Lots of Link (TRILL) and Shortest Path Bridging (SPB) are some protocols already being used by some companies and are expected to be widely used in the future as the alternatives to varieties of Spanning Tree Protocols. The great question is how these new protocols will be addressing the massive growth in Network Virtualization and cloud computing.

With the introduction, use, and future prospects of these protocols, some questions are sure to hit the mind of networking personnel, for example, what is behind these new technologies, how do they work and how do the various alternative protocols differ from each other?

In the following chapters, this thesis goes through the basic ideas on Spanning Tree Protocols, their limitations and challenges, and elaborates particularly on the TRILL and SPB protocols. It explains the operation of these protocols, their reliability and examines their possible implementation in the future.

2. SPANNING TREE PROTOCOL

The Spanning Tree Protocol (STP) is an OSI layer 2 protocol that is required to create a loop free network topology. With its mechanism of having redundant links, it also helps avoiding the chances of MAC address instability. The person behind the invention of spanning tree protocol is Radia Perlman. She designed the spanning tree algorithm, which creates a network topology in a tree shape, bypasses the probable loop, and makes the network stable and redundant.

When there is more than one path in the network, and STP is not enabled, then the network loop is created. A layer 2 loop causes various problems, such as MAC address instability, broadcast storms and multiple frame transmission in the network.

MAC Address Instability: One of the problem caused in the looped environment of the network is that the MAC table is constantly updated and it rises with the flooding of the broadcasts resulting the network failure.

Broadcast Storms: When a network loop occurs, each switch may flood broadcasts continuously without stoppage. This problem is referred to as Broadcast Storms.

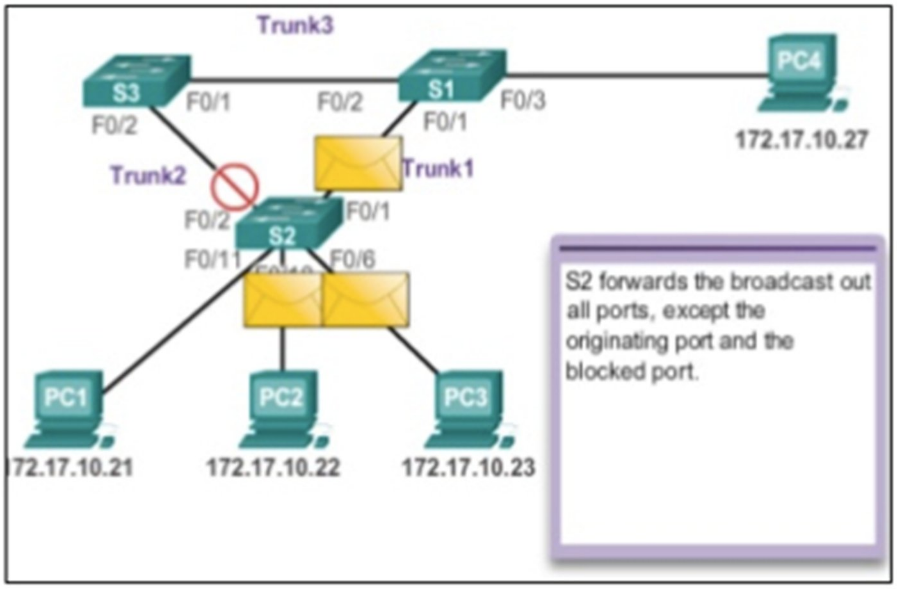
Multiple Frame Transmission: Multiple copies of the same frame are delivered to the destination computer causing unrecoverable errors as the protocol expects to receive only a single copy of the unicast frame. [1]

2.1. STP Operation

STP uses three main concepts, namely root bridge, port roles, and path costs to find out the link to be used in a redundant network topology.

Redundancy is always very important in a bridged network as it secures the availability of the network during the single link failure because of the faulty network device or cable. Because of the physical redundancy in the network design, loops and duplicate frames occur. To avoid the severe consequences of the loop and duplicate frames in the network, the Spanning Tree protocol was introduced.

STP creates only one logical path between all the destinations by intentionally blocking the used ports and other redundant path which may cause network failure. During the network failure, to compensate them for a network cable or a switch failure, STP recalculates and unblocks the required ports allowing the redundant path and hence ensuring redundancy.



S2 forwards the broadcast out all ports, except the originating port and the blocked port.

Figure 1. STP Operation. [6]

Thus, the network switches running the STP are able to compensate the network failure by strategically or dynamically or automatically unblocking the previously blocked ports and providing an alternate route for the traffic by permitting them. Here the point of our discussion is to original spanning tree, so to avoid the confusion, the term IEEE 802.1d is used.

2.1.1. Spanning Tree Algorithm: Port Roles

For the network redundancy, IEE802.1d uses the Spanning tree algorithm to find out which redundant path or say the ports in the switch are to be blocked. Among all the switches, one switch is referred by STA as Root Bridge, which is used later as the referencing point to calculate the alternate path during the network failure. The root bridge is chosen by the election among all the switches in the network topology. While the switches participate in exchanging the BPDU frames, the switch with the lowest bridge id (BID) is determined and is automatically assigned to be the Root bridge.

A BPDU is a messaging frame exchanged by Switches for STP. BPDU contains BID which gives the information about the switch sending the BPDU. BID contains a priority value, MAC address of the source and an optional extended system ID. All these three are calculated to determine the lowest BID value, hence assigning the root bridge.

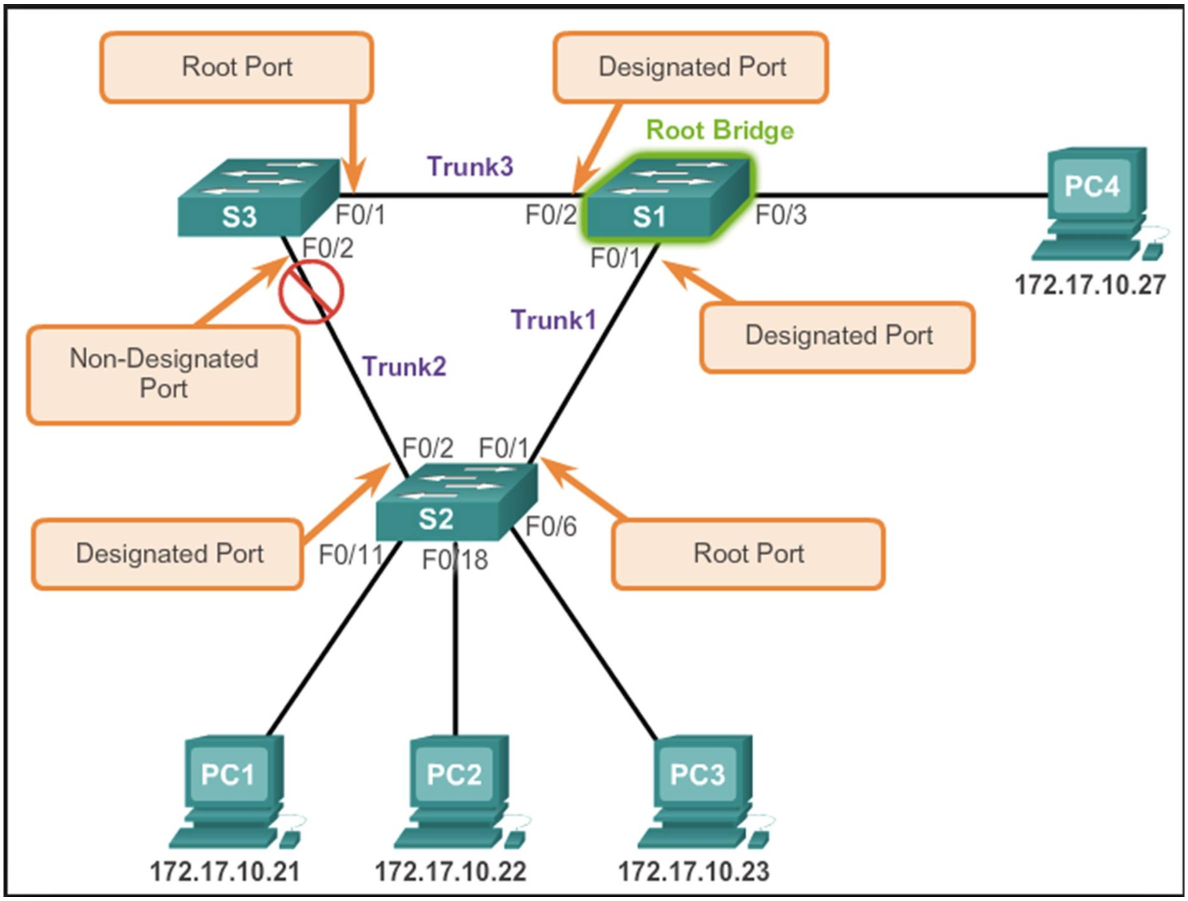


Figure 2. Port roles [6]

Root Ports: They are the ports in the switch closest to the root bridge.

Designated ports: The designated ports are determined on a per-trunk basis. In a trunk line if one end is a root port, then the other end ports become the designated ports. In the root bridge, all ports are designated ports.

Alternate or Backup Ports: These are the ports which are blocked by the STP to ensure the loop-free topology. When any network failure occurs, to compensate the failed switch or a network cable, this port is unblocked and becomes operational hence ensuring redundancy.

Disabled Ports: The switch port which is shut down is a disabled port.

2.1.2. Spanning Tree Algorithm: Root Bridge

A switch is determined as a root bridge in every Spanning tree instance which works as a reference point for all the calculations in determining root ports, designated ports, alternate ports and disabled ports.

2.1.3. Spanning Tree Algorithm: Path Cost

When the root bridge is determined, the Spanning Tree Algorithm selects the best path to the root bridge from all the destination in broadcast domain. The path information is calculated by adding each port costs and the root bridge. The speed of the Ethernet ports determines the default port costs. The following figure shows the best path to the root bridge:

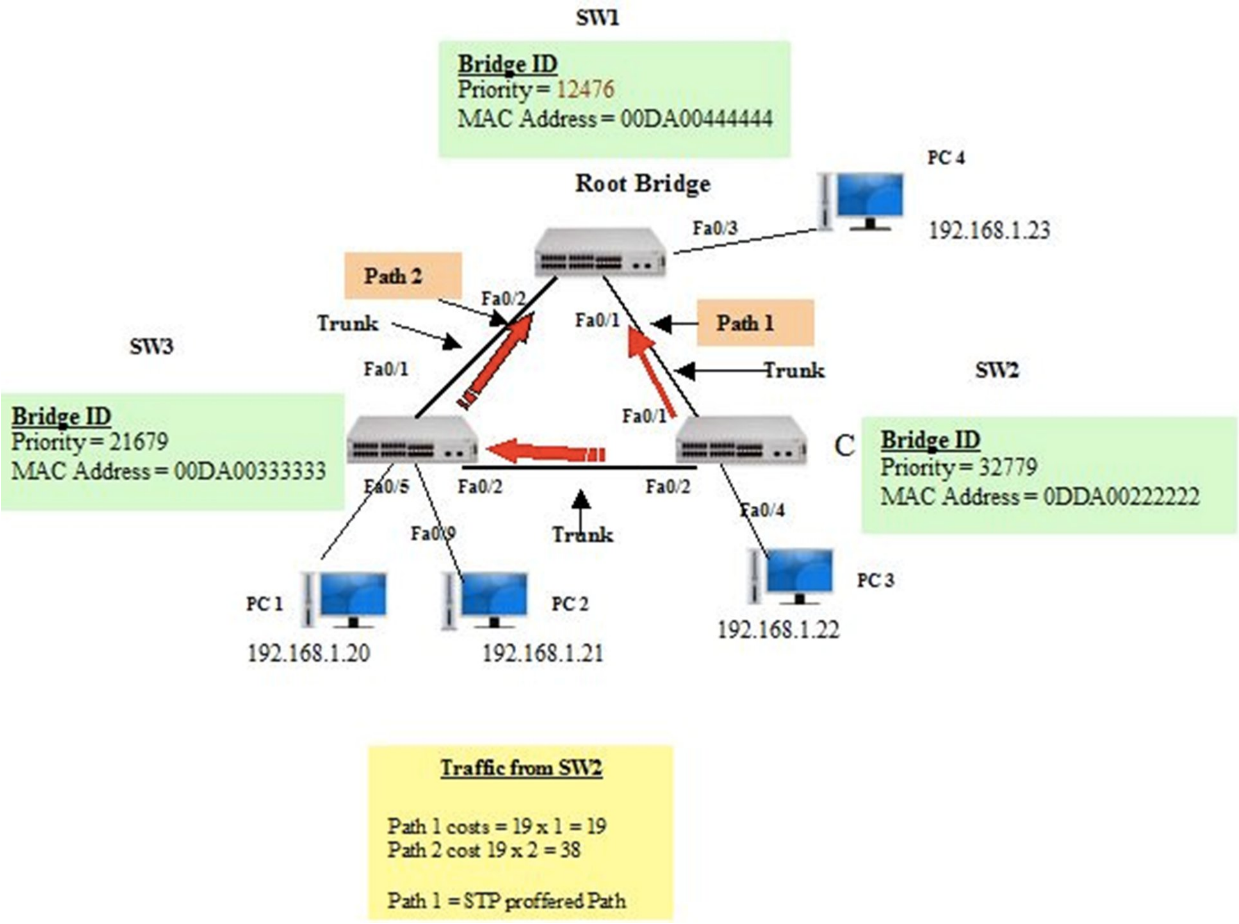
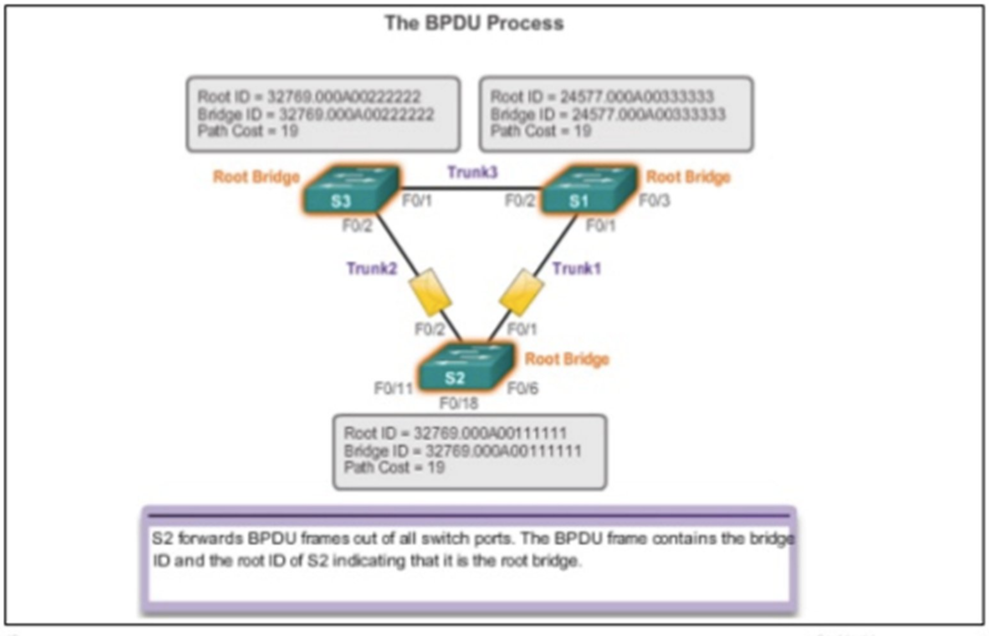


Figure 3. Path to the root bridge. [6]

Although default port costs are already defined, each port cost can be configured to new value and it can be in between 1 and 200,000,000.

2.2. BPDU Frame

The exchange of BPDUs frame plays an important role to determine Root Bridge. Figure 4 shows the 12 independent fields of BPDUs which explains the path and priority required while determining the root bridge and the path to the same.



The BPDU Process

Root ID = 32769.000A00222222
Bridge ID = 32769.000A00222222
Path Cost = 19

Root ID = 24577.000A00333333
Bridge ID = 24577.000A00333333
Path Cost = 19

Root ID = 32769.000A00111111
Bridge ID = 32769.000A00111111
Path Cost = 19

S2 forwards BPDU frames out of all switch ports. The BPDU frame contains the bridge ID and the root ID of S2 indicating that it is the root bridge.

Figure 4. BPDU propagation and process. [6]

2.3. Extended System ID

The use of the extended system ID includes VLAN ID in the BPDU frame along with the Bridge priority enhancing the Spanning Tree Protocol in order to support the VLANs. Normally, the value of bridge priority ranges between 0 and 65535 having the default value 32768 but with the extended system ID, the value is the sum of VLAN ID and the bridge priority value making it range between 0 to 61440. The Extended System is depicted in the following Figure 5.

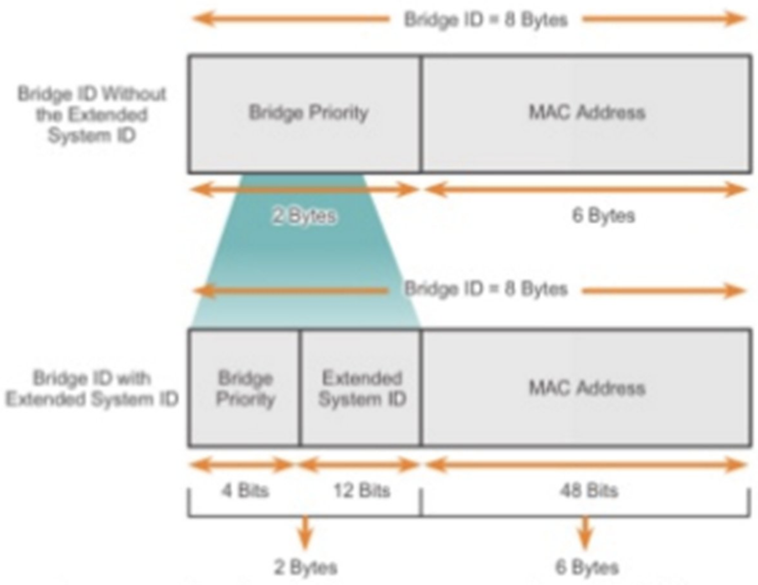


Figure 5. Extended system ID. [6]

3. TYPES OF SPANNING TREE PROTOCOL

Spanning Tree Protocols and their advancement gradually with the time are explained briefly below.

3.1. STP

The original version was earlier introduced as 802.1d-1998. It works as single SPT instance in the whole bridge network regardless of the number of VLANs. The CPU and memory required in this version of protocol are very low because of its single instance in the entire network. In addition, the single root bridge and tree are present. Since the network traffic flows over the same path for all VLANs, this is outdated and the version is slow in network convergence.

3.2. PVST+

This enhancement of the STP provides a separate path for the traffic flow for each different VLAN configured in the network. This separate instance supports PortFast, UplinkFast and BackboneFast, BPDU guard, BPDU filter, root guard, and loop guard. More CPU and memory is required but it allows per-VLAN root bridges. The Spanning Tree can be optimized for the traffic in each VLAN.

3.3. 8021d-2004

8021d-2004 is an updated version of STP standard incorporating IEEE802.1w

3.4. RSTP (IEEE802.1w)

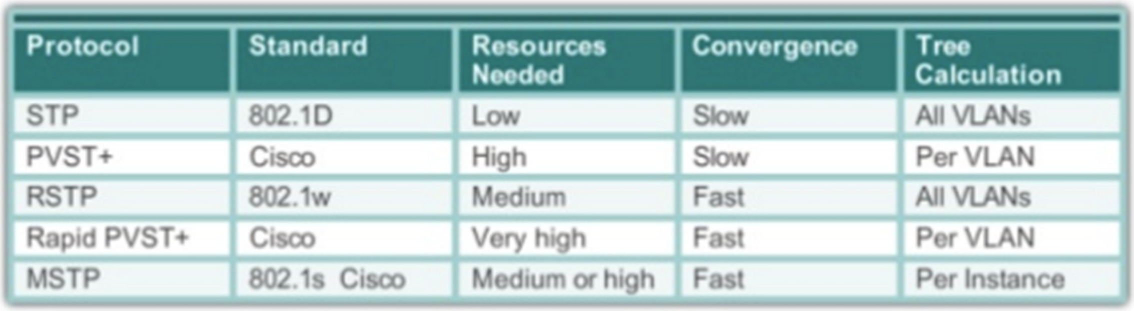
This is the enhanced version of the standard STP which provides faster convergence. Although this version is updated with many convergence issues, it provides a single instance STP, not addressing suboptimal traffic flow.

3.5. RPVST+

This is a Cisco enhancement of RSTP which uses PVST+ providing separate instance of RSTP per VLAN. This version supports all PortFast, UplinkFast, BackboneFast, BPDU guard, BPDU filter, root guard, loop guard and also addresses the convergence issues and suboptimal traffic flow. However, this version requires larger CPU and high memory.

3.6. MSTP

This maps multiple VLANs into the same Spanning Tree instance. It addresses the high memory and CPU requirement issues of Rapid PVST+ [1].



Protocol	Standard	Resources Needed	Convergence	Tree Calculation
STP	802.1D	Low	Slow	All VLANs
PVST+	Cisco	High	Slow	Per VLAN
RSTP	802.1w	Medium	Fast	All VLANs
Rapid PVST+	Cisco	Very high	Fast	Per VLAN
MSTP	802.1s Cisco	Medium or high	Fast	Per Instance

Figure 6. Spanning tree protocol characteristics. [6]

4. LIMITATIONS AND CHALLENGES OF SPANNING TREE

The basic function of the Spanning Tree Protocol, which operates at layer 2 of the OSI model is to cut loops and avoid redundant links created in the bridge networks. STP elects the ports which forward or block the traffic with the help of exchange of BPDUs between the bridges. This protocol, when failed, is very difficult to troubleshoot as it depends mostly on network design. Therefore, the most important part of the troubleshooting is better carried out before the problem is encountered.

4.1. Spanning Tree Failure

Failure in STP mostly directs to loop bridging. It is wrong to assume that the spanning tree failure occurs because of some bug. The bridging loop in STP always occurs because of blocked ports forwarding the data or vice-versa.

4.1.1. Spanning Tree Convergence

When the Spanning tree converges, it results in excessive loss of BPDUs and because of this, a block port goes into a forwarding state which eventually leads to STA failure.

There are many different situations in the bridge network that result in STA failure while most of the failure cases are related to excessive loss of BPDUs.

4.1.2. Duplex Mismatch

When the duplex mode is set to full on one side of the point-to-point network and the other side is in auto negotiation mode, it results in half duplex. This mismatch is one of the most frequently occurring configuration errors. A bridging loop easily occurs and the scenario becomes worse when a switch that sends BPDUs is set to half duplex and its connecting peer on the other end has full duplex mode as shown in Figure 7.

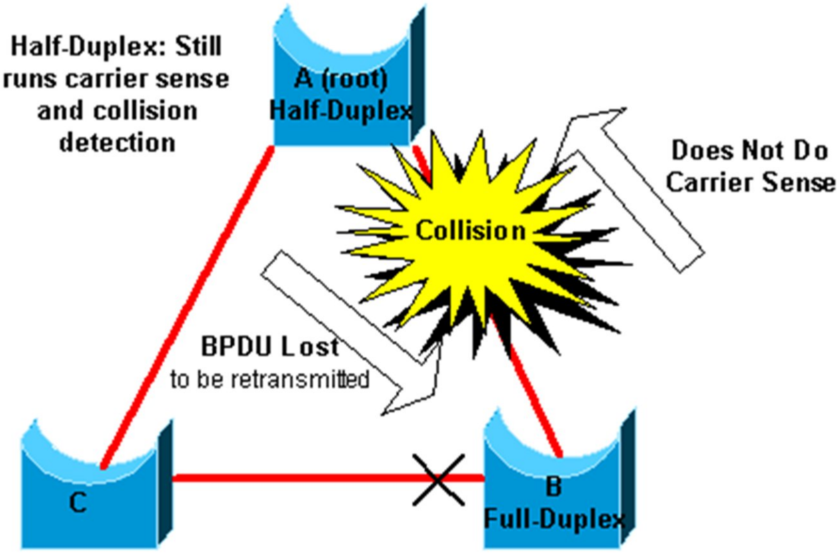


Figure 7. Duplex mismatch [2].

The problem grows larger when the duplex mode of the Bridge’s port sending the BPDUs is assigned as half duplex but the duplex mode of the peer port connecting to the other side of the same link is full-duplex. This is shown in Figure 7 where the outgoing port of the Bridge A is set to half-duplex mode whereas the incoming port of the bridge B has full-duplex mode set. This duplex mismatch simply results in the occurrence of a bridging loop. Because of the full duplex mode, before the link access, carrier sense is not seen in Bridge B. It begins to send frames without caring that Bridge A already using the link. Then the situation leads to a problem in Bridge A as it starts detecting the collision. Immediately, Bridge A runs the back off algorithm. When there are more and more packets coming from B to A, all the packets sent by A including the BPDUs collided and are eventually dropped. Now, since the bridge B does not receive any BPDU information from A, the root bridge becomes unreachable. After that, when B opens its port connected with C, the occurrence of a loop is observed.

4.1.3. Unidirectional Link

One of the frequent causes of loops in the bridge network is the Unidirectional link. Generally, the problem that occurs in fiber links without detection and failure in the transceiver usually causes Unidirectional links. Any Unidirectional link which provides one-way transmission of the traffic is harmful to the STP operation. Figure 8 is an example of this problem:

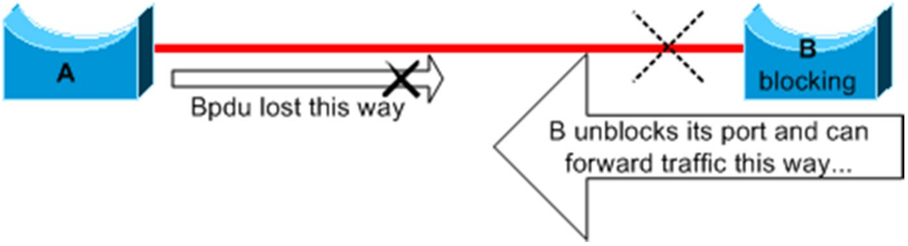


Figure 8. Unidirectional link. [2]

The unidirectional link in the figure above drops the transmission from switch A to B but the traffic is transmitted the other way around. Thus, all the BPDUs transmitting from A to B are lost while the transmission from B is active as it unblocks its port which eventually creates a network loop. Hence, when there is this kind of failure in startup, STP does not converge properly. Reboot may help with the problem temporarily in the case of duplex mismatch but has no effect at all in this case.

The UDLD protocol designed by Cisco detects the unidirectional link before the forwarding loop is created on layer 2 and accordingly breaks the loops by disabling the required ports. So running the UDLD in the bridged environment helps resolving the problem.

4.1.4. Packet Corruption

Incorrect cable or its length, duplex mismatch and so on are the reasons behind Packet Corruption which results in a similar kind of failure in the network. This failure often leads blocking ports to the forwarding state which eventually creates problem in the STA algorithm and its functioning.

4.1.5. Resource Errors

Sometimes the recourses seem to be insufficient, when STP is running in software and when the CPU of the switch is over utilized, it creates problems in the normal transmission of BPDUs resulting in STP failure.

4.1.6. PortFast Configuration Error

PortFast is usually enabled on a port connected to a host and this port transitions into forwarding state skipping the first stages of the STA when the link comes up. PortFast feature is never used on switch interface that connects to other switches, hubs or routers because this may create a network loop.

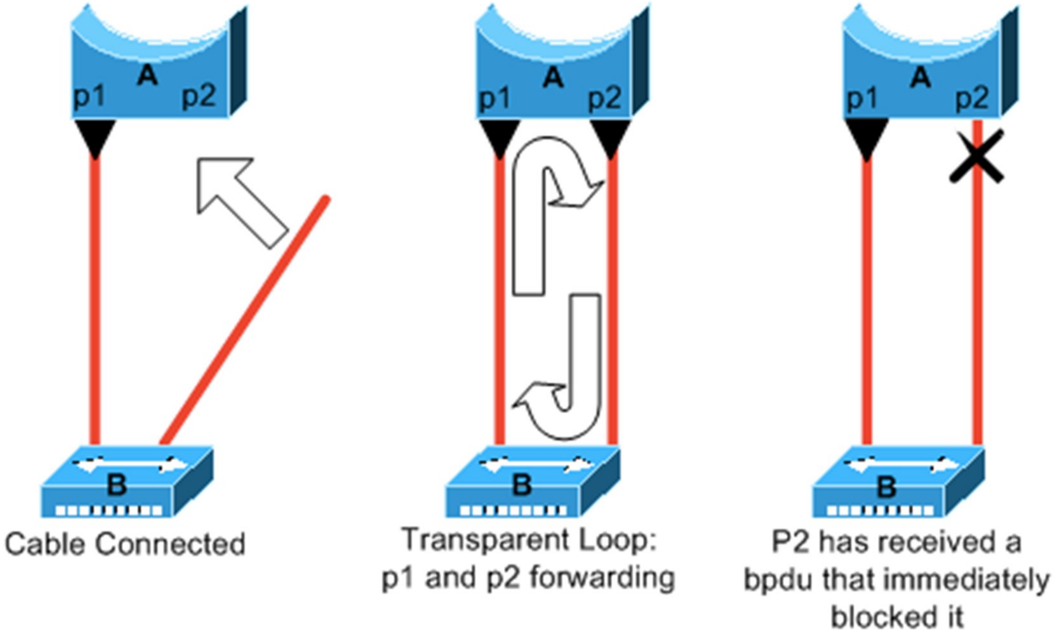


Figure 9. PortFast configuration error. [2]

In Figure 9, device A is a Bridge with ports P1 and P2, and device B is a hub. In A, port P1 is in forwarding state and port P2 is configured with a PortFast. When the second cable is connected to A, P2 immediately goes into forwarding mode resulting in the loop between P1 and P2. When the BPDU reaches P1 or P2, either of these ports go to blocking state and hence the loop is stopped. This kind of transient loop, however, has some issues. In case of very heavy looped traffic, there may not be a successful transmission of BPDU resulting in no stoppage of the loop or delay in stopping the loops. Thus, this problem creates further problems in network convergence and may bring down the network in severe cases.

Although the PortFast is configured, STP is still active in the port or the interface. When a switch has a lower bridge priority in comparison to the current root, the bridge is connected to a port having the PortFast configuration and the latter could be elected as the root bridge. This kind of re-election of root bridge can create an adverse situation in the network. In order to solve this problem, most switches running either CatOS or Cisco IOS software have a unique feature called BPDU guard. The BPDU guard helps solving the problem by disabling the port configured with PortFast when it receives a BPDU.

4.1.7. Awkward STP Parameter Tune and Diameter Issues

When setting the value for the max-age parameter and the forward delay, one should be very cautious. An overestimated value on these makes the STP topology unstable, causing a loop to appear in the network. At the same time, the diameter of the bridged network is really important to be considered as the distance between two distinct bridges should never be more than seven hops away between each other. This is because the default value of the maximum network diameter imposed by STP parameters is seven. One should pay special attention before changing the STP timers from the default value. Trying to converge the network faster with aggressive values may create a dangerous situation in the network resulting in poor stability of the STP.

4.1.8. Software Errors

Some software errors cause STP to fail. For example, EtherChannel in some specific cases causes STP failure. Generally, software errors occur because of numerous different factors, although they can not be described adequately, software errors can be minimized by not ignoring typical BPDU transmission and transition of the ports in Bridged networks. [2]

4.2. Need for Simplifying Network Virtualization

In the present world talking about the context of data center, shifting to server virtualization is immense and par above anyone could think. Virtualization is extremely beneficial as it saves a great deal of costs for hardware, has a lot of space, has fewer number of servers used and it reduces the power and cooling problems in the network. The traditional networking system is directly affected by this virtualization of servers. Spanning Tree Protocols and their advancements, namely RSTP and MSTP, are not enough to handle the loops and to create redundant network topology. This shifting to Virtual Machines adds other requirements in networks regarding the extensions of the Layer 2 VLANs between the data centers in different locations or between different sections within the data center. New configurations or changes in configurations within the layer 2 are definitely required and use of a non-optical path for the traffic between the data centers may be seen in many instances.

Network virtualization has become prominent for all around success for many large companies as they have started building their own cloud computing environments. The realization of the essence of cloud computing and its benefits, such as easiness in adding the resources and services as required, accessing the applications in different locations anytime, have been felt by the companies. So the need of establishing a virtualized data center backbone have become essential. The following figure explains this need. The five infrastructures, namely Ethernet Scalability, Always-On Resiliency, Cohesive Management, Energy Efficiency and Workload Mobility are the results and benefits of network virtualization between Data Centers.

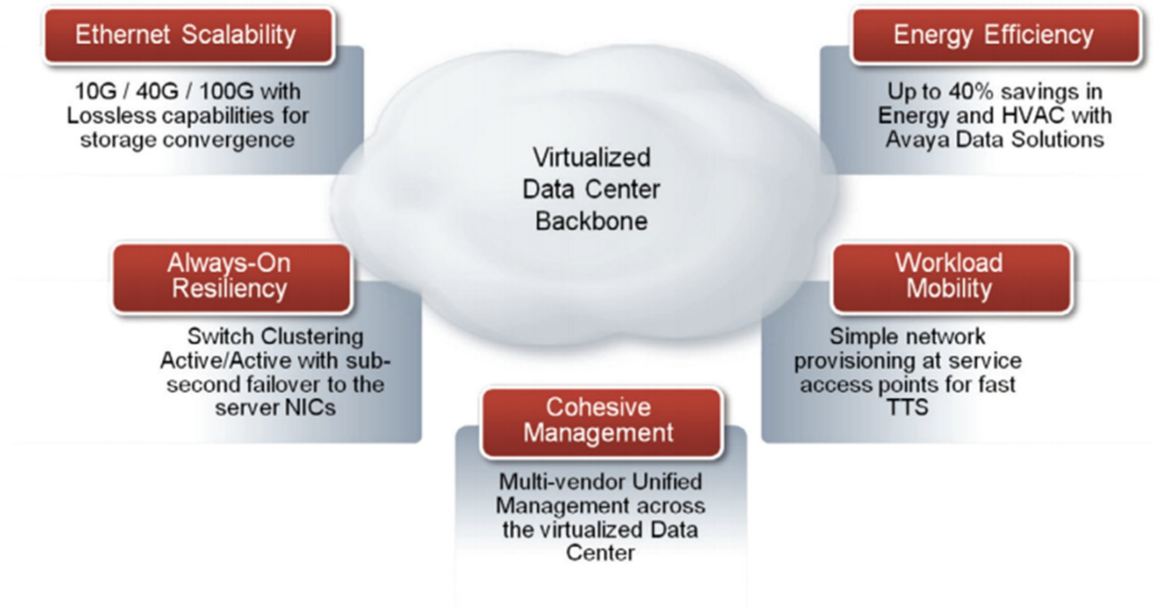

Ethernet Scalability
10G / 40G / 100G with Lossless capabilities for storage convergence
Energy Efficiency
Up to 40% savings in Energy and HVAC with Avaya Data Solutions
Virtualized Data Center Backbone
Always-On Resiliency
Switch Clustering Active/Active with sub-second failover to the server NICs
Workload Mobility
Simple network provisioning at service access points for fast TTS
Cohesive Management
Multi-vendor Unified Management across the virtualized Data Center

Figure 10. Network virtualization in data center. [14]

To handle the essential obligations of the data center in terms of cloud computing, Virtual Machine mobility, advance control in terms of traffic flow, right and correct use of bandwidth, and reduction in terms of number of devices used, Network Virtualization is most necessary. Shortest Path Bridging (SPB) and Transparent Interconnections of Lots of Links (TRILL) aim to achieve the same without creating any complexity in the network. Both of these emerging protocols' goal is to create a strong and efficient network topology eliminating the STPs by supporting both multipath forwarding and providing an easy solution to failure unlike Spanning Tree Protocols. SPB and TRILL promise the same and have, hence, emerged as unprecedented solution to STP. In the next chapter both these protocols are explained in detail. [14]

5. ALTERNATIVES TO SPANNING TREE PROTOCOLS

Spanning Tree Protocol and its advancements RSTP and MSTP, are not enough to handle loops for the Ethernet networks in present and future world where network virtualization has become most important for the overall success of the networking industries. Now, when the companies have started to have their own cloud computing environment, these protocols are not enough to handle the network redundancy as they end up with a strict tree topology and are inefficient to deal with the scalability and reliability issues.

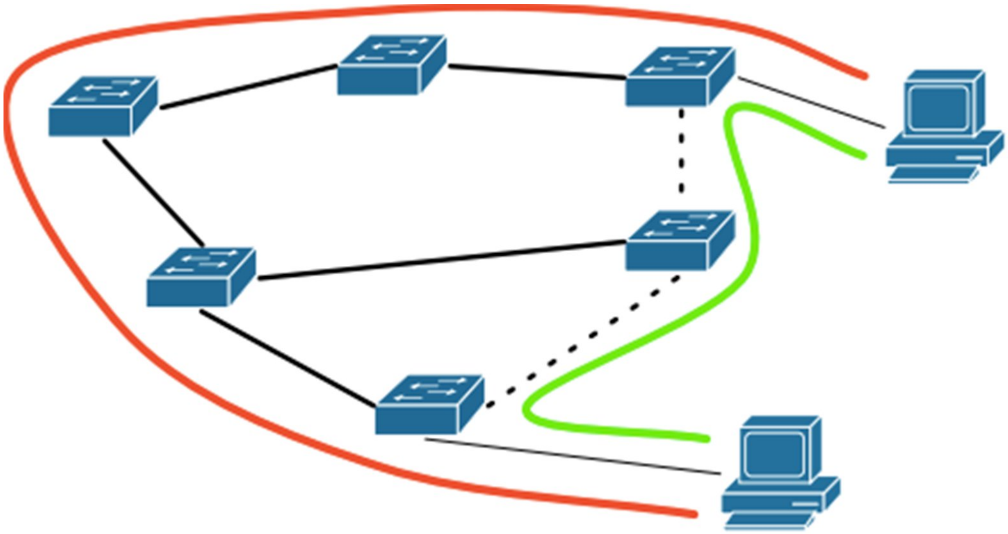


Figure 11. Limitations on Spanning Tree Protocol. [15]

In Figure 11, the Spanning Tree Protocols does not forward the traffic through the shortest path as indicated by the green line. It is restricted in the tree topology formed and results in flowing of the traffic to the switch in the left and then to the destination switch through the longer path as indicated by the red line. This restricted topology here also does not allow the alternative path between the two nodes because if that happens, loop in the network topology will be created. In addition, the converging time of the spanning tree protocol has been a serious issue. It takes several seconds of time to converge in case the topology changes. It is stated that the original STP takes about 30 seconds to converge in case of a difficult network scenario. Thus, with the STP protocols, it has been a great challenge for the network managers to forecast the active forwarding link. Although the protocol runs with the definite algorithm, the end result is uncertain and unpredictable with the use of STP, RSTP or MSTP in large network environment. [15]

With the limitations in Spanning tree protocols, the two different protocols namely: TRILL and SPB have emerged as the most promising alternatives to STPs for the future. TRILL that functions within IEEE 802.1-complaint Ethernet broadcast domain uses IS-IS routing to distribute link state information and determine shortest path over the network. It is a pure routing protocol which does not require an IP for transferring data.

The IEEE 802.1aq standard Shortest Path Bridging (SPB) brings into the use of Shortest Path Trees (SPTs) as an alternative to the trees being used by STP, RSTP or MSTP. The SPTs used in this protocol always ensure that the traffic in the network is transferred through the shortest path possible between the two bridges. [8, 11]

These two protocols, their basic concepts and background, functioning, similarities and differences are explained in details in the following sections.

5.1. Transparent Interconnection of Lots of Links (TRILL)

TRILL is a layer 2 Ethernet protocol which operates being based on IS-IS routing. It distributes the link state traffic replacing STP and calculates the shortest path through the bridged network. It uses IS-IS routing because IS-IS is a layer 2 routing protocol which does not need an IP for the transmission of data. In-between the routing bridges, Trill data packets and IS-IS routing packets are exchanged. Through IS-IS hello frames, Routing Bridges discover other bridges automatically and, hence, any other difficult configuration is not required. MAC-addresses of the end devices are known at the end-points of the path (ingress and egress) of a Trill domain only. In the process, the core bridges do not require to learn the MAC-address information of the end devices. Trill is mainly required in mesh topologies especially in big data centers where STP is not enough. [3, 5]

Trill, proposed and discovered by the original inventor of STP, Radia Perlman, is in process of being standardized. About Trill and its development, the official working group (WG) states its work on the same as follow:

“The TRILL WG has specified a solution for shortest-path frame routing in multi-hop IEEE 802.1-compliant Ethernet networks with arbitrary topologies, using an existing link-state routing protocol technology and encapsulation with a hop count. The current work of the working group is around operational and deployment support for the protocol. This includes a MIB module and other pieces needed for operations, but also additional ways to extend and optimize TRILL for the properties of the networks on which it is deployed". [7]

The functioning of Trill is best explained in the figure below.

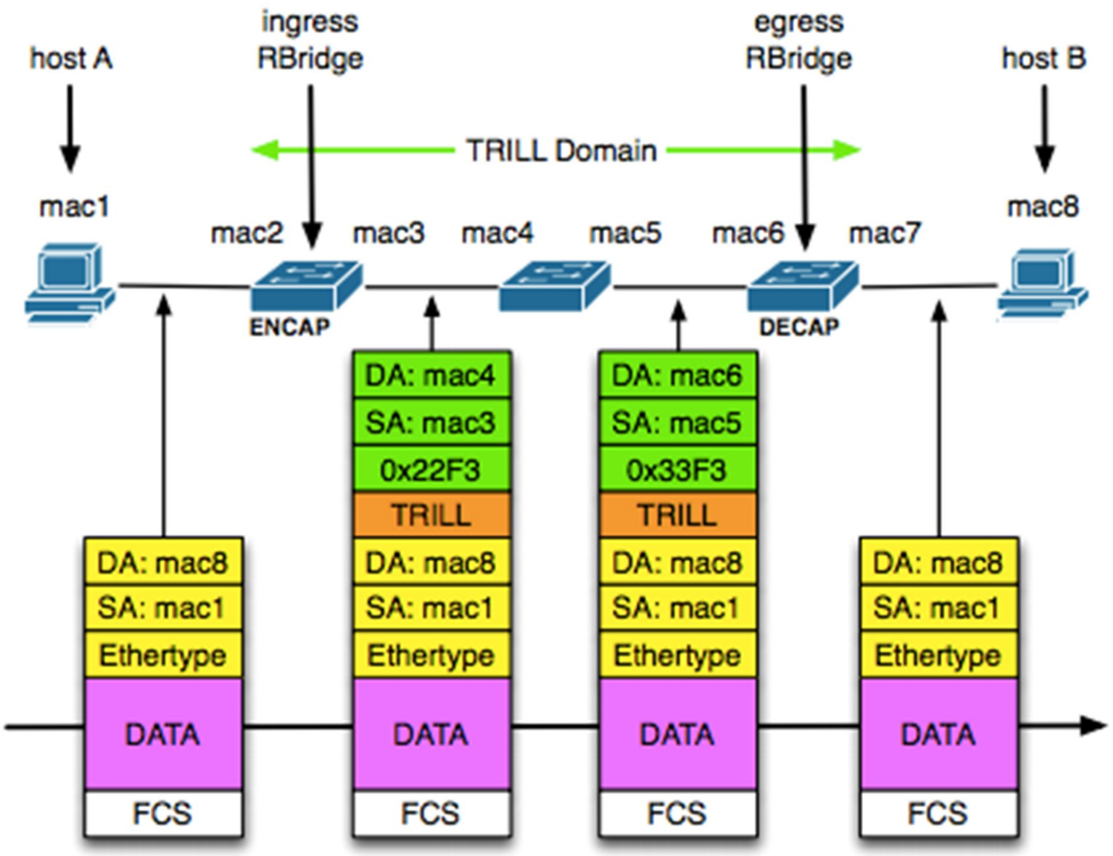


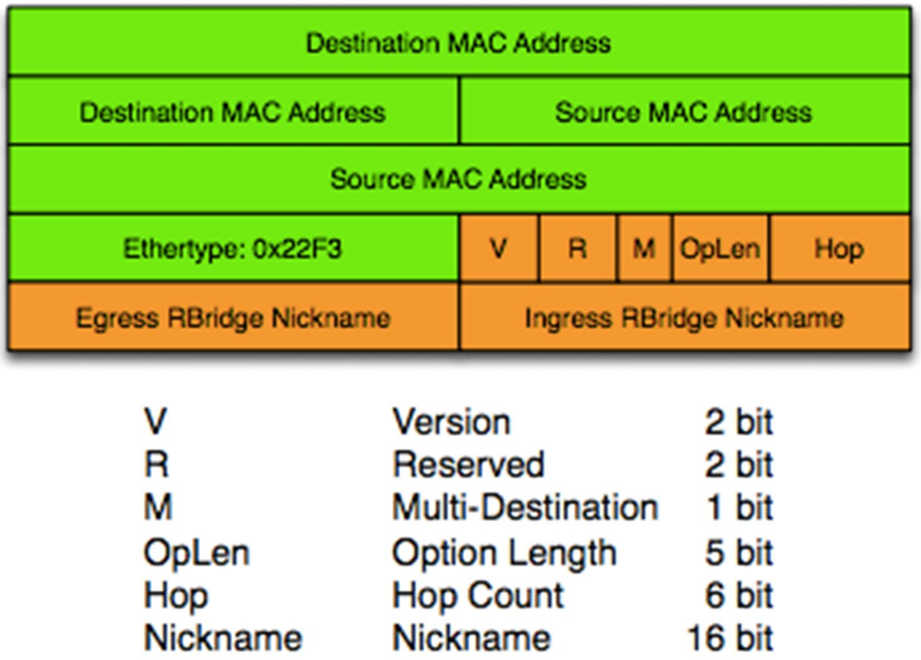
Figure 12. Encapsulation and Decapsulation in Trill. [15]

The Ethernet frames are encapsulated by the RBridges at the incoming edge(Ingress) and those frames are routed through the whole TRILL Domain using ISIS routing and the frames are decapsulated again at the outgoing Edge Bridge (at egress), as shown in the above figure

Host A sends an Ethernet frame which is destined to Host B, the frame when reaches the first RBridge at ingress, it is encapsulated with a TRILL header and an Ethernet header by the RBridge. The TRILL header contains hop count and of 2-byte identifying information of both ingress Rbridge and egress RBridge. In the process, Egress RBridges is chosen in the process through the shortest path possible to reach the destination host B. The Ethernet header is re-written at each RBridge starting at the ingress edge and in each hops in the path determined by

IS-IS routing. The Source address is the MAC address of the outgoing interface of the RBridge and the destination address is the MAC address of the next hop RBridge. When the frames pass each bridge, the hop count is gradually decreasing from the TRILL header. When the encapsulated frame reaches the RBridge at egress, the encapsulation headers are taken off and the original Ethernet frame is delivered to the destination host B.

Except for the RBridges at Ingres and Egress, the bridges in between could be any traditional switch and in any number (many in case of big topology). These in-between switches do not necessarily need to support TRILL as their task is to flow the traffic with the information of MAC address and VLAN ID, if present. As mentioned above, TRILL uses trill header and Ethernet header for the propagation of the Traffic. Trill header and Ethernet header can be seen in the figure below.



V	Version	2 bit
R	Reserved	2 bit
M	Multi-Destination	1 bit
OpLen	Option Length	5 bit
Hop	Hop Count	6 bit
Nickname	Nickname	16 bit

Figure 13. Ethernet and Trill Headers. [15]

Adapting 0x22F3 as Ether type, Trill Protocol can encapsulate both untagged and tagged customer frames and while encapsulating tagged frames the tag in the frame is always preserved. Unicast frames are forwarded through the shortest path between the ingress and egress while the multi destination frames are sent using the distribution trees. In the TRILL domain, there are one or multiple distribution trees and the ones which might me used are precompiled by the RBridge. In the trill domain, there is one RBridge, chosen by all the RBridges, that decides the number of trees, the types of trees used and the tree number of each tree for all the RBridges [15].

5.2. Shortest Path Bridging (SPB)

SPB is an open and standard based solution to STPs, RSTPs and MSTPs which is highly reliable and scalable in multipath broad network environment where the services are usually controlled at the edges [10]. It was standardized and approved as IEEE 802.1aq standard on March 28, 2012. As an alternative to spanning trees used by STP, RSTP and MSTP, SPB uses shortest path trees (SPTs) which ensures the shortest path between the two bridges. The area where SPTs are used may exist side-by-side with the area where STPs, RSTPs and MSTPs are used. Within the SPB region, a Common and Internal Spanning Tree (CIST) is used as a default spanning tree and the IS-IS link state routing protocol, ISIS-SPB, is used in order to transfer information within the bridges to calculate shortest path trees. All the source bridges in the SPB region to all remaining bridges in the same region are involved in calculating Shortest Path Trees. All the bridges involved in calculating the Shortest path trees calculate exactly the same set of trees which is defined as the SPT algorithm. In SPTs, both unicast and multicast traffic take the same path and also both outgoing and incoming traffic move through the same path making SPTs bidirectional. Multiple Equal Cost Trees (ECTs) are calculated in order to support Load Balancing. In total 16 ECTs use 16 different SPT algorithm tie-breakers and, hence, each ECTs uses different tie-breakers. All the SPTs sharing the same ECTs tie-breaker and supporting one or multiple VLANs within the region are defined as SPTs set. The VLANs are distributed over up to 16 SPTs set while load balancing which means per packet load balancing between two end devices is absent unless the end devices use different VLANs for different packets flow. The SPB is divided into two different categories: Shortest Path Bridging VID (SPBV) and Shortest Path Bridging MAC (SPBM). [9, 15]

5.2.1. Shortest Path Bridging VID- SPBV

The Shortest Path Bridging with the Vlan Identifier is what is known an SPBV. All the VLANs that are managed by SPBV use an SPT set. Either manually or automatically, SPVID is allocated to every SPT individually in the set. ISIS-SPB is used to send the mapping information from SPVID to SPT to the other bridges. The VIDs which are in a C-TAG or S-TAG of a customer frame is converted to SPVID in accordance with the SPT which accepts that VID at the incoming region of SPBV. In the outgoing area SPVID is converted back to the primary VID. In case the customer frames have no C-TAG or S-TAG, a tag with SPVID is added by SPBV at Input region. The tag is taken away again at the output. MAC addresses of the end devices are known at each individual bridge during the path.

5.2.2. Shortest path bridging MAC- SPBM

SPBM, where M stands for MAC(Media Access Control) is used with PBB(Provider Backbone Bridges). As the customer frames in PBB are encapsulated with an ethernet header, PBB is often called MAC-in-MAC. The BEB(Backbone Edge Bridge) adds a PBB header at input device and removes it at the output. The information contained in the PBB header are source MAC address of the incoming BEB, destination MAC address of the outgoing BEB and 24 bit I-SID(Backbone Service Instance Identifier). Because of 24 bit I-SID, 224 different services can be practised which can be said as similar to Virtual Private Network(VPN). The frames are carried only between the ports that map to the same ISID. Each individual I-SID is mapped to B-VID although mapping multiple I-SID to the same B-VID is also possible. The frames are transported within the PBB backbone in accordance with the destination B-MAC address and B-VID. Customer VIDs are separated by the PBB from backbone VIDs and only at the edges of PBB region end station MAC learning takes place. A single B-VID is used for all SPTs in a set by SPBM and the PBB backbone bridges identifies only B-MAC.

The SPBM encapsulation and decapsulation which are carried out by BEB B and BEB E respectively are shown in Figure 14. Many different service interface options are provided for the customer ports Through PBB BEB the customer frames can be both tagged or untagged with a C-TAG. All the required frames are mapped to I-SID and are taken forward with no S-TAG. The service interface which is deployed or transporting the frames forward with the S-TAG is called S-tagged Service Interface, also known as Q-in-Q frames. Two different types of S-tagged service interfaces can come across; in the first one, one-to-one mapping is done between S-TAG and I-SID whereas in the second one multiple S-TAGs are mapped to the same I-SID. The first case shows S-TAG not being carried within the PBB backbone and the second case shows S-TAG being carried in the PBB header so that at output BEB would know about the particular S-TAG the original frame contained.

In the figure 14, in S-tagged Service Interface on BEB B, various S-tags are mapped on I-SID 100 expalin the S-tag being transferred through PBB backbone. Both frames sent to end station H2 from H1 are not tagged and both are attached to S1 and S2 respectively. The switch port is configured in PVID(Port based VID) mode and the VID value is set to 20 to which the end station devices are connected. In the service provider switches BEB B and BEB E, both the switches S1 and S2 establish the connection with an S-tagged interface. The user frame is sent from H1 by S1 to the BEB B as an S-Tagged frame.

The S-Tag which is to be used for each VLAN has to be configured on S1. Either all or each C-VLAN could be directed to the same or different S-TAG. In order to carry various customer VLANs through the same S-TAG over the same I-SID service instance, the same S-TAG has to be used. In figure 14, we can see that BEB B using an S-Tagged Service Interface and is encapsulating the delevered frame with a PBB header. To determine the particular provider service to be used in the backbone, a mapping between S-tag and I-SID occurs at BEB B. The mapping can be one-to-one or bundled with various S-TAGs on the same I-SID. With the second case, S-TAGs must be included in the PBB header so that the particular S-TAG to be used in the de-capsulated frame is identified by the decapsulating side. Here, in the figure C-VID is mapped to I-SID 100 in the end. The shortest path is used to forward the encapsulated frame to the output BEB E through the SPBM region with the information of the MAC address as the destination address in the PBB header. Here in the PBB header, the VID used is the B-VID that is meant to contain I-SID 100. The administrator configures the the mapping in between I-SID and B-VID either one-to one or many-to-one. The frame is finally decapsulated at the output by BEB E and this decapsulated S-Tagged frame moves to S2, where S-TAG and C-TAG are removed and the resulting frame is sent to H2(End Station). In order not to let customer see S-Tag and to provide direct C-VID to I-SID mapping, vendors may decide to have the functionality of S1 and BEB B within the same box. Thus, a simple V-LAN mapping is carried out between provider services and VLANs benifiting the customer as the particular provider services can be chosen by deploying a particular VLAN. BEBs come to identify the end stations around in the decapsulating pricess and the MAC learning is possible in the same process in the SPBM region. [15]

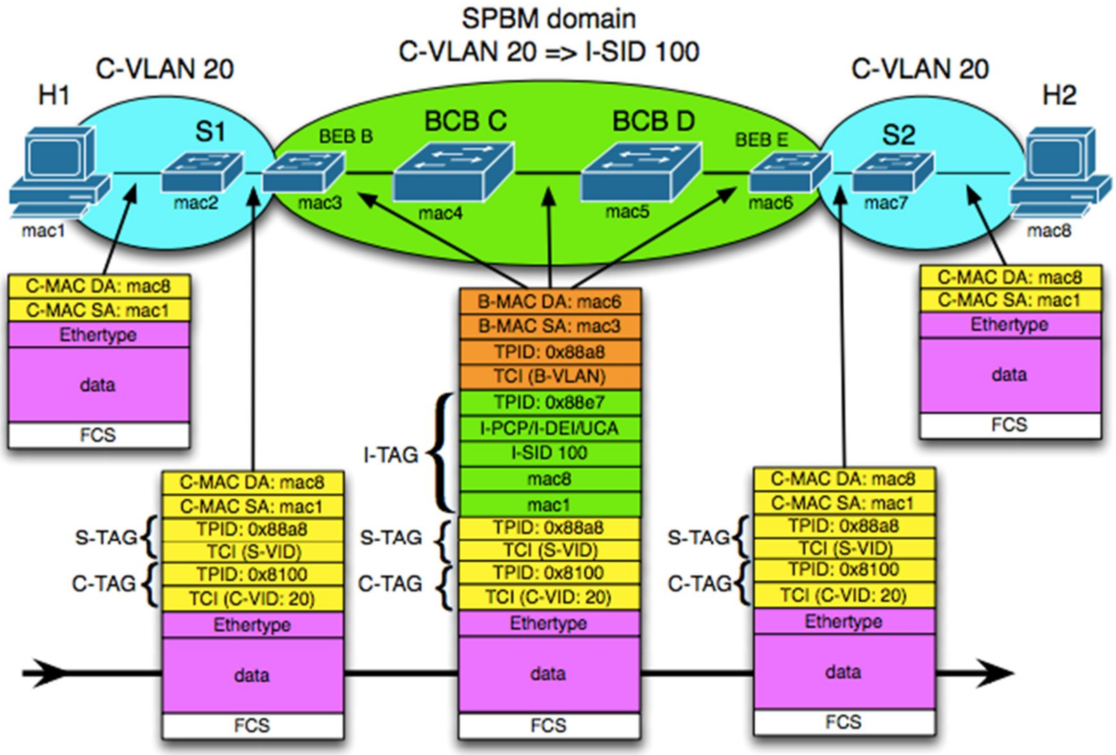


Figure 14: Mapping of C-VLAN 20 over I-SID 100 and S-Tag bundled Interface In SPBM. [15]

5.3. Comparison of TRILL and SPB

There are few similarities among TRILL and SPB technologies which are listed below:

- Both of these technologies seek to establish a powerful layer 2 topology by excluding the Spanning Tree and adopting multipath forwarding with localized failure resolution.
- Both of them support multi-pathing and interoperability within spanning tree.
- IS-IS is used in both technologies as a layer 2 routing protocol with low touch configuration.
- Cut-through switching is possible in both but it is difficult in TRILL due to options field in header.
- Both the technology dynamically changes network paths for traffic flows and the unicast traffic path is the shortest path based on IS-IS calculations.

However, several features exist among these technologies which make them differ from one another. The differences are explained in Figures 15 and 16, and Table 1 on the basis of background and technology.

Table 1. Differences between SPB and TRILL. [14]

	SPB	**TRILL**
Background and basics	It was developed by Nortel as a Provider Link State Bridging (PLSB) to the IEEE.	It was introduced in 2006 as a IETF proposed standard to the IEEE.
	The support for SPB technology along with participation in IEEE SPB initiative has been announced by Avaya, Alcatel, Hauwei and Cisco.	The intention to support for TRILL along with the participation in the IETF TRILL initiative has been announced by Cisco, Brocade and Juniper.
	It is an established technology for many years in the carrier market.	It is a new technology without any former roots.
Technology	It uses RPFC for loop prevention.	TTL(Time to Live) is used to minimize the loop and support the formation of non-congruent trees along with RPFC.
	It uses the former IS-IS with TLV extensions.	It creates a new type of IS-IS instance with new PDU types
	Virtualization support by service instance using I-SID (16 Mio)	It can support up to 4000 VLANs.
	Election processes is pre-provisioned	Election processes are Designated Forwarder, Root Bridge, IS-IS nicknames per RBridge
	It supports MAC-in-MAC encapsulation.	It supports TRILL header encapsulation.
	The multicast traffic path is between two end nodes, which is similar to unicast and bi-directionally congruent and tree based on source node.	The multicast traffic path locate on selected root bridge unicast which can have totally different multicast traffic path (which can lead to out-of-sequence packets while transforming from BR/MC path to unicast path).

	The output processing for multicast is not needed.	The output processing for multicast is needed due to MAC header change output port.
	SPB	**TRILL**
	The customer MAC learning is packet-based at the end of SPB network.	The customer MAC learning is packet-based at the end access port and ESADI protocol.
	The out of sequence packet is not possible.	The out of sequence packet is possible when transformation of Dest MAC occurs from unknown MAC to known.
	The service is aggregated for example multiple VLANs could be mapped as Service Instance.	There is no service aggregation.
	The traffic is assigned to the head end through the shortest path by using link based metrics to calculate path.	The shortest path is assigned for unicast along with Layer 2 header swap in every RBridge and uses link based metrics to calculate path.
	The troubleshooting is easier as entire path can be seen through the network and has IEEE/ITU based Ethernet OAM tools.	Traffic needs to be inspected by hop-by-hop basis to understand the path. There is no OAM tools.
	New hardware is not required as 802.1ah, 802.1ad, 802.1ag is supported in many platforms.	The absence of OA&M hardware support requires the new hardware.
	IP/SPB Draft supports the extension of layer 3 and IP VPN	There is no integration of layer 3 and IP VPN extensions.

The lookup and forwarding comparison between TRILL and SPB is shown in figure 15. In the TRILL, since the hop-by-hop method is used for forwarding, every node look ups of TRILL header with MAC swaps, the decrement of TTL and the recalculation of Frame Check occur. Thus, the difficulty arising in the network and troubleshooting becomes problematic. No simple measure is identified and, hence, hop-by-hop should be used to manage troubleshooting. However, the difficulty is eliminated with the implementation of SPB as it uses a simple MAC forwarding table look up and designate traffic to a shortest path to the required output point. The troubleshooting is, hence, resolved as the whole flow can be determined using source and destination address. [14]

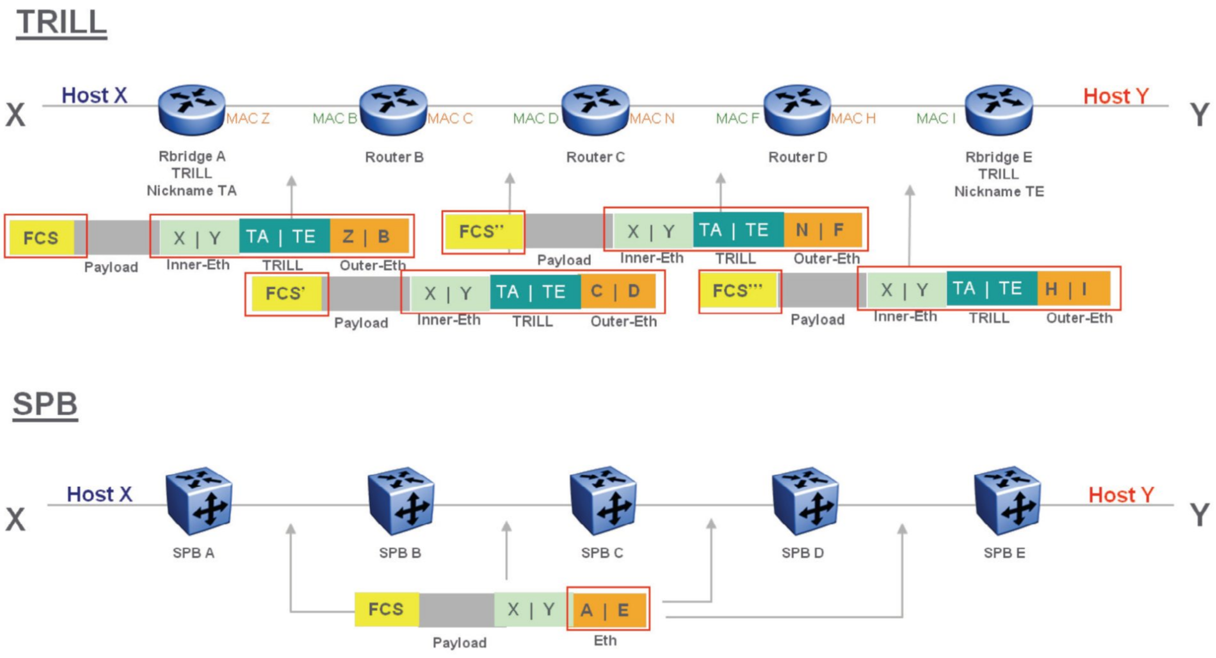
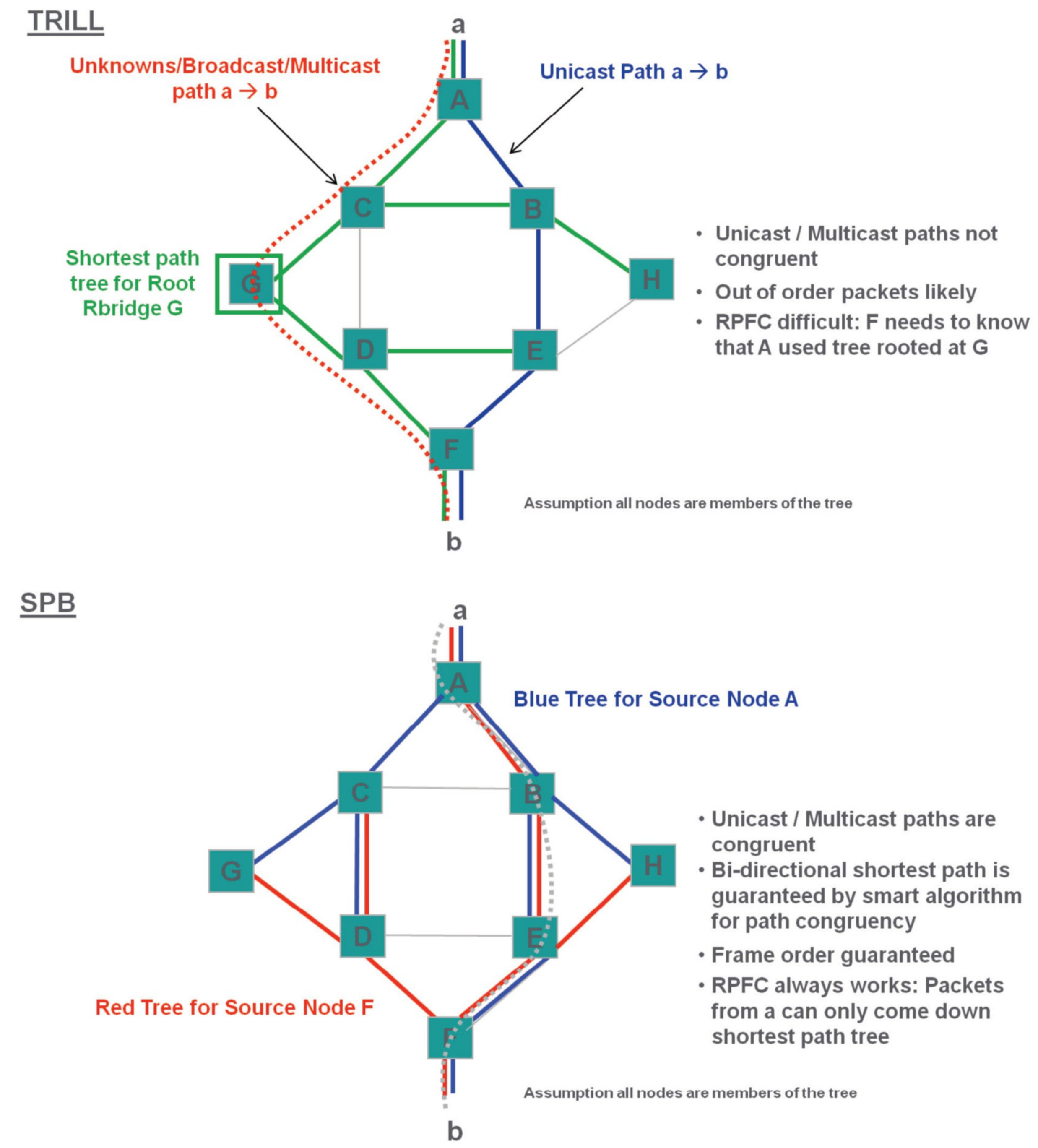


Figure 15: TRILL and SPB lookup/forwarding. [14]

Figure X shows the shortest path tree for RBridge G. The broadcast/multicast path of TRILL is shown in the figure by the red dotted line from a to b which are not consistent. Out of order packets are likely to be seen and F must have the knowledge that A used tree rooted at G. In case of SPB, the unicast/ multicast path is consistent and the shortest path is guaranteed in two directions by a smart algorithm. Here, the frame order is guaranteed and RPFC is always reliable.


TRILL
a
Unknowns/Broadcast/Multicast path a → b
Unicast Path a → b
A
C
B
Shortest path tree for Root Rbridge G
G
H
D
E
F
b
SPB
a
A
Blue Tree for Source Node A
C
B
G
H
D
E
Red Tree for Source Node F
F
b

- Unicast / Multicast paths not congruent
- Out of order packets likely
- RPFC difficult: F needs to know that A used tree rooted at G

Assumption all nodes are members of the tree

- Unicast / Multicast paths are congruent
- Bi-directional shortest path is guaranteed by smart algorithm for path congruency
- Frame order guaranteed
- RPFC always works: Packets from a can only come down shortest path tree

Assumption all nodes are members of the tree

Figure 16: Shortest Path Trees. [14]

6. FUTURE ASPIRATIONS AND DEVELOPMENT

The limitations of the spanning tree protocol have become obvious in recent years. The demand for smooth networking in big mesh topology, its scalability and reliability are not met by the STPs. Many large companies nowadays have their own cloud computing environments and because of that, network virtualization has become very important for them. With this development in networking industries, the need for suitable extension for layer 2 topology has become most important. It has been a challenge for companies to deploy a suitable protocol in place of long used trusted Spanning Tree Protocols. IETF TRILL and SPB have emerged as the possible reliable solution.

Both TRILL and SPB using the IS-IS routing are developed to implement robust forwarding in big meshed topologies. Because of the use of these protocols, networks links are handled in a more efficient way making the network more resilient than when using the Spanning Tree Protocols. SPB or TRILL-based cloud computing is much better than STP-based. The main advantage of using either TRILL or SPB is in case of failure both the protocols are easy to support and the networks are developed to easily handle the end devices.

Having said this, IETF TRILL has a broad market value and is being used by the majority of the companies. The reason for it is also because TRILL are proposed and developed by Cisco. SPB, on the other hand, is supported by relatively fewer companies and is mostly backed by Avaya. Having two variants for the use in case of VLANs and MAC, namely SPBV and SPBM, SPB is a very strong competitor for TRILL. Although TRILL is a newer technology than SPB, it has a wider support which shows that it has a somewhat brighter future in networking industries.

7. CONCLUSION

The alternative protocols to be used in place of traditional Spanning Tree protocol have been very a hot topic in the networking industries. For companies, this is extremely important as it deals with the use of right protocols for the future, meeting the future customer requirements and the overall success in scaling up the networks towards the coming times. To investigate the future alternatives to overcome the shortcomings of the current Spanning Tree technologies has been a matter of great interest for the author.

IETF TRILL and SPB are the two protocols that seem to offer the best options for Local Area Network redundancy in the future. The scalability issues, convergence issues and unpredictability, and many more problems of the traditional STPs are met by these new technologies. With the use of shortest path, both SPB and TRILL are sure to make the LAN networking strong, redundant, and highly scalable in the future.

SPB seems to be a step forward compared to traditional standards. SPB standardized by both IEEE and IETF implements all the advanced ideas regarding MPLS and, naturally, it has backward compatibility with spanning tree.

TRILL, on the other hand, is backed by the large vendors, such as Cisco and Brocade, and does not need an IP for the transmission of data. It works on the basis of transferring the IS-IS hello frame through which RBridges discover other RBridges automatically and hence avoiding any other difficult configuration in the core.

Both standards, IETF’s TRILL and the IEEE802.1aq SPB, are somewhat similar as both try to solve the same problem in the network. Using IS-IS and working with VLANs, the goal of both protocol is to provide an easy, transparent, automatic, and yet simple solution to the vendors.

REFERENCES

1. Lewis, W. 2008. *LAN Switching and Wireless CCNA Exploration Companion Guide*. 1st ed. United States of America: Cisco Systems, Inc.

2. Cisco Systems. 2015. *Spanning Tree Protocol Problems and Related Design Considerations*. [ONLINE] Available at: http://www.cisco.com/c/en/us/support/docs/lan-switching/spanning-tree-protocol/10556-16.html. [Accessed 11 November 2015].

3. TechTarget. 2011. *Transparent Interconnection of Lots of Links (TRILL)*. [ONLINE] Available at: http://searchnetworking.techtarget.com/definition/Transparent-Interconnection-of-Lots-of-Links-TRILL. [Accessed 8 November 2015].

4. Cisco Systems. 2010. *Catalyst 2960 and 2960-S Switch Software Configuration Guide: Cisco IOS Release 12.2(53) SE1*. [ONLINE] Available at: http://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst2960/software/release/12-2_53_se/configuration/guide/2960scg.pdf. [Accessed 5 January 2016].

5. Touch, J. & Perlman, R. 2009. *Transparent Interconnection of Lots of Links (TRILL): Problem and Applicability Statement*. [ONLINE] Available at: http://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst2960/software/release/12-2_53_se/configuration/guide/2960scg.pdf. [Accessed 12 December 2015].

6. Cisco Networking Academy. 2015. *Lan Redundancy*. [ONLINE] Available at: http://www.slideshare.net/vuzlego/chapter2-lan-redundancy. [Accessed 25 February 2016].

7. Cisco Systems. 2011. *Introduction to TRILL*. [ONLINE] Available at: http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_14-3/143_trill.html. [Accessed 8 December 2015].

8. Extreme Networks. 2016. *TRILL vs. SPB*. [ONLINE] Available at: http://www.trex.fi/2014/xtrm-trill-vs-spb.pdf. [Accessed 15 February 2016].

9. Fedyk, D., Ashwood-Smith, P., Allan, D., Bragg, N.,& Unbehagen, P. 2012. *IS-IS*

Extensions Supporting IEEE 802.1aq Shortest Path Bridging. [ONLINE] Available at: https://tools.ietf.org/pdf/rfc6329.pdf. [Accessed 15 February 2016].

10. The Networking Nerd. 2013. *Avaya and the Magic of SPB*. [ONLINE] Available at: https://networkingnerd.net/2013/10/14/avaya-and-the-magic-of-spb/. [Accessed 12 December 2015].

11. Fratto, M. Network Communications. 2010. *Shortest Path Bridging Will Rock Your World*. [ONLINE] Available at: http://www.networkcomputing.com/networking/shortest-path-bridging-will-rock-your-world/1689548769/page/0/1. [Accessed 5 February 2016].

12. Ashwood-Smith, P. 2010. *Shortest Path Bridging IEEE 802.1aq Tutorial and Demo*. [ONLINE] Available at: https://www.nanog.org/meetings/nanog50/presentations/Sunday/IEEE_8021aqShortest_Path.pdf. [Accessed 23 March 2016].

13. Kerner, S.M. 2012. *Will TRILL or Shortest Path Bridging Win Out? Enterprise Networking Planet*. [ONLINE] Available at: http://www.enterprisenetworkingplanet.com/datacenter/will-trill-or-shortest-path-bridging-win-out.html. [Accessed 15 December 2015].

14. Avaya. 2010. *Compare and Contrast SPB and TRILL*. [ONLINE] Available at: http://techdata.com/business/avaya/DataCenterSolutions/files/A%20-%20Why%20Avaya/2%20-%20Learn%20More%20About%20VENA/SPB-TRILL_Compare_Contrast-DN4634.pdf. [Accessed 19 December 2015].

15. Van Der Pol, R. 2012. *TRILL and IEEE 802.1aq Overview*. [ONLINE] Available at: https://kirk.rvdp.org/publications/TRILL-SPB.pdf. [Accessed 10 January 2016].