



Leopold-Franzens-Universität Innsbruck

Institute of Computer Science
Information Security

**Visualizing network topologies with the help of the
Spanning Tree Protocol**
Bachelor Thesis

Alexander Schlögl

advised by
Dr. Pascal Schöttle

Innsbruck, September 20, 2016

Inhaltsverzeichnis

List of Figures	i
List of Listings	ii
1 Introduction & Motivation	2
2 Background	5
2.1 Existing alternatives	5
2.2 Spanning Tree Protocol (STP)	5
2.3 Technologies used	9
3 STP Tree Generator	12
3.1 Class Structure	14
3.2 Packet Handling	15
3.3 Communication	19
3.4 Uncertain Assumptions	21
3.5 Combining the data	24
3.6 Visualization	28
3.7 Installation & Usage	31
4 Software Switch Testing Utility	34
4.1 Saving Switching Data	34
4.2 Handling STP Packets and Port States	35
4.3 Handling Non-STP Packets	37
4.4 Installation and usage	39
5 Testing	40
5.1 Setup	40
5.2 Tests	41
6 Conclusion	49
Bibliography	51

Abbildungsverzeichnis

1.1	An example of a broadcast storm	3
1.2	Broadcast propagation in an STP network	3
2.1	An example of STP port states	6
2.2	An STP BPDU	6
3.1	Information gained on STP build up	13
3.2	Information lost during tree build up	13
3.3	An example case of false assumptions about network structure	22
3.4	Assumed and actual message ages	23
3.5	An example for STPViz output	28
5.1	The expected physical and logical network topologies	41
5.2	Expected output compared to STPViz output (Plug and PlayTest)	43
5.3	The result of the Tree Establishment Test	44
5.4	Expected output compared to STPViz output (Tree Removal Test)	45
5.5	The physical topology during the Slow Dynamic Test	46
5.6	The packet capture not showing a TC flag	47
5.7	Expected output compared to STPViz output (Slow/Fast Dynamic Change Test)	48

Listings

2.1	Pcap Callback Prototype	9
2.2	JSON Example	10
2.3	A TikZ Example	10
3.1	The SpanningTree header file	14
3.2	Filtering for STP packets	15
3.3	Going Through the Payload	16
3.4	Checking for Previously Known Information	17
3.5	Bridge Data Update	18
3.6	Client-Server Transmission	19
3.7	Server-Parser Transmission	20
3.8	Removing Incorrect Bridge Data	22
3.9	Combining the Bridge Data	24
3.10	Recursive Tree Combination Algorithm	24
3.11	Adding subtrees in the place where they belong	26
3.12	Comparing SpanningTree objects for equality	27
3.13	.tex code for Figure 3.5	28
3.14	TikZ conversion function, part 1	28
3.15	TikZ conversion function, part 2	29
3.16	Tikz conversion function, part 3	29
3.17	Tikz conversion function, part 4	30
3.18	Calculating the maximum width of a tree	31
4.1	The Containers for Switching Data	34
4.2	Updating Port States	35
4.3	Handling Non-STP Packets	37

Abstract

Monitoring the logical topology of Spanning Tree Protocol (STP) enabled networks is a difficult task, and often only accomplished by using other network management protocols. Not all networking hardware supports these protocols, making this a suboptimal approach. In this thesis we present STPViz, a tool which can monitor nodes in a network using only passively obtained STP information. It uses a distributed client-server architecture to gather STP packets at multiple points in the network. STP transmits only local data, but STPViz can still use this distributed information to draw conclusions about connections in the network. As our method of assuming connections in the network does not guarantee correctness, we provide a simple and computationally cheap way of eliminating errors. The error eliminating accuracy increases with the number of STPViz instances in the network, going up to 100% if there is an STPViz instance connected to every networking device. Due to inconsistencies we found in STP implementations during testing, we had to implement the protocol ourselves to be able to test STPViz. STPViz passed all the tests we designed. It can handle node discovery, connection discovery, device removal and even dynamic topology changes. These test results leave us with the impression that with usability improvements STPViz can become a useful addition to the network monitoring suite for environments where the networking hardware does not support more complex management protocols.

Zusammenfassung

Die Überprüfung von Netzwerktopologien, welche das Spanning Tree Protocol (STP) verwenden, ist eine schwere Aufgabe, welche oft nur mit Hilfe von Netzwerkverwaltungsprotokollen bewältigt werden kann. Da nicht alle Netzwerkgeräte diese Protokolle unterstützen, ist dies eine suboptimale Lösung. In dieser Arbeit präsentieren wir STPViz, ein Werkzeug, welches Netzwerktopologien nur anhand von passiv erhaltenen STP Paketen beobachten kann. Es verwendet eine verteilte Client-Server-Architektur um STP Pakete an mehreren Orten im Netzwerk zu erhalten. Obwohl STP nur lokale Daten überträgt, kann STPViz durch diese verteilten Informationen Vermutungen zu Verbindungen im Netzwerk anstellen. Da unsere Annahmen nicht immer korrekt sind, beschreiben wir auch eine einfache und leicht berechenbare Methode zur Entfernung dieser Fehler. Die Genauigkeit unserer Methode wächst mit den STPViz Instanzen im Netzwerk, und erreicht 100% wenn eine Instanz mit jedem Netzwerkgerät verbunden ist. Durch Ungenauigkeiten in STP Implementationen, die wir beim Test von STPViz fanden, waren wir gezwungen, STP selbst zu implementieren, um STPViz zu testen. STPViz bestand alle unsere Tests. Dabei wurden seine Fähigkeiten zur Entdeckung von Geräten, Annahmen zu Verbindungen zwischen Geräten, die Entfernung von Geräten sowie dynamische Topologieveränderungen getestet. Diese Testresultate verleiten uns zur Annahme, dass STPViz mit verbesserter grafischer Ausgabe ein nützliches Werkzeug zur Netzwerkkontrolle in Umgebungen ohne Unterstützung von komplexeren Netzwerkverwaltungsprotokollen sein kann.

Kapitel 1

Introduction & Motivation

As networks grow larger and more complex, redundancy becomes desirable and necessary. Adding redundant links always introduces loops to the network. Switches and hubs respond to broadcast messages by sending them over every port (except for the one the packet was received on). This is done in order to ensure every node in the network receives the message. In conjunction with loops, this behaviour leads to broadcast messages being repeated indefinitely. With the right network setup, these messages can also multiply exponentially. Conditions like this are called broadcast storms[19]. Figure 1.1 shows how a broadcast storm forms. In Figure 1.1a you can see how Bridge *A* starts the broadcast. Bridges *B* and *C* then propagate this broadcast over every other port. This leads to the broadcast being sent in both directions between bridges *B* and *C*, as well as twice along to bridge *D*, as seen in Figure 1.1b. As bridges do not keep messages in memory after they have been sent, they do not recognize broadcast messages they received before. Thus, they simply propagate the messages in compliance with their policy, leading to the packets multiplying (Figure 1.1c). After the next iteration of propagations the same broadcast message will be sent in both directions over every bridge connection, shown in Figure 1.1d.

To prevent this the Spanning Tree Protocol (STP)[15] was created. It creates an overlay network, in the shape of a tree, to remove loops from the network. In this overlay, broadcast storms are not possible anymore. An example for broadcast propagation in an STP overlay network is shown in Figure 1.2. The overlay ignores the connection between bridges *C* and *D*. This leads to the broadcast being propagated to every node in just two steps, without any extra packets being generated.

For this thesis, we will only talk about switches. Hubs, while still existent, are hardly ever used and not STP capable. We will also refer to switches as bridges, in order to be compliant with STP nomenclature[15].

Networks with STP utilization are usually quite large. Additionally, changes and outages are hidden by STP and redundant links in the network. This can make it hard to keep track of the current network layout. Debugging STP configurations is also a difficult and time consuming task. It is possible to check STP configurations via the Simple Network Management Protocol (SNMP). However, this requires an administrator to connect to every single bridge and collect the STP information manually. While there are tools to automate this, they rely primarily on SNMP, which may not be supported by all bridges in the network.

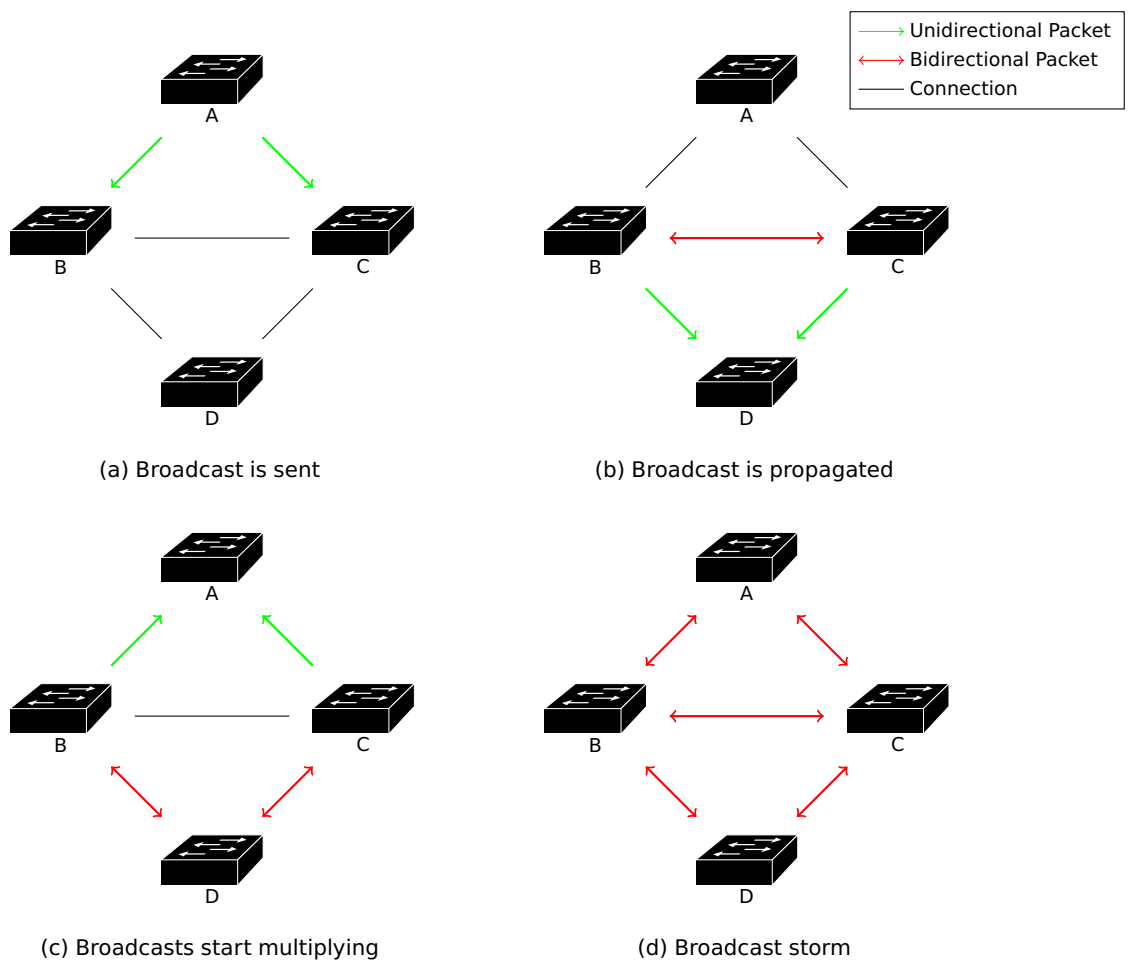


Abbildung 1.1: An example of a broadcast storm

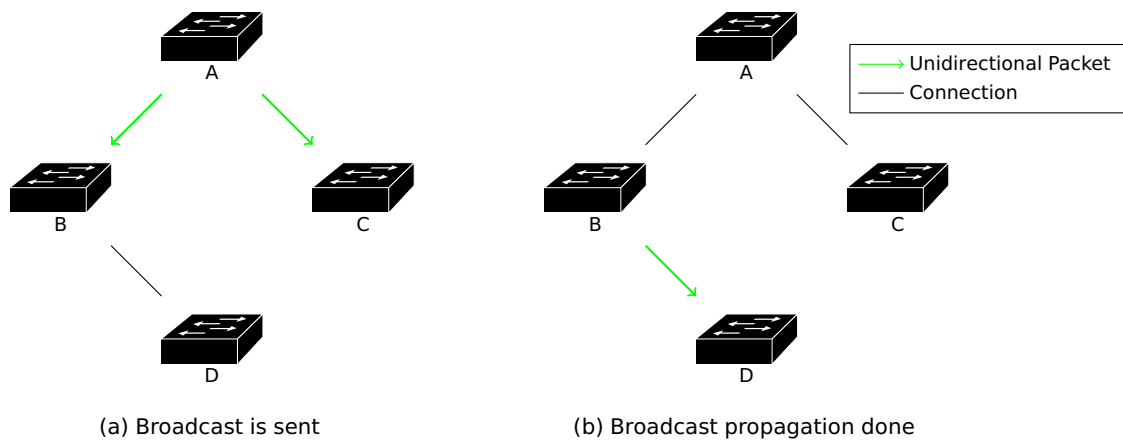


Abbildung 1.2: Broadcast propagation in an STP network

The aim of this thesis is to create a versatile tool for STP visualization. We chose to name this tool *STPViz*. STPViz has the following requirements:

- **STP only:** STPViz should only rely on STP packets
- **Low traffic:** STPViz should create as low a traffic as possible
- **Passive:** STPViz should not provoke any changes in the network and only use packets obtained passively
- **Distributed:** STPViz is to be run on multiple nodes to send information to a central server where everything is pieced together
- **Low hardware load:** STPViz should run on simplistic hardware (e.g. a Raspberry Pi) or on a regular PC without slowing it down noticeably
- **No maintenance:** once set up, STPViz should not require any maintenance (except when updating to a new version)
- **Multiple trees:** if the clients send data from multiple different spanning trees (e.g. one tree per VLAN), these need to be kept separate.

The remainder of this thesis is structured as follows:

Chapter 2 will cover the background knowledge needed to fully understand STPViz we developed. It includes information on STP, how it works, the packets it uses, the Spanning Tree Algorithm and extensions to the protocol. Libraries and tools we used for this thesis are also explained.

In chapter 3 we will discuss the implementation of STPViz. We will cover our methods of saving the data, client-server communication, and packet handling. Discovering bridges and connections between them, as well as our ways to mitigate errors that may occur during discovery are explained in great detail. It also describes how we obtain the final visualization. A short explanation of the installation requirements and command line parameters is included as well.

Chapter 4 shows the details of our *software-switch* tool, which we had to develop due to problems we encountered with the STP implementations in free router OSs[14][5]. It shows how we handle STP and non-STP packets, in addition to our approach to port states. Again a short section with instructions is included.

Chapter 5 explains how and to what extent we tested STPViz, as well as the results. Here we describe the setup for each of our five tests and the steps required to perform them. We show a graphical representation of the expected result for every test, and compare it to the output of STPViz. Additionally, we analyse the results, and give insight into what happened and why.

Finally, Chapter 6 will summarize this thesis and give an outlook on possible future improvements to STPViz.

Kapitel 2

Background

2.1 Existing alternatives

As of writing this thesis, only very few tools support monitoring of STP topologies. The ones we found are LorientPro[11], LiveAction[10] and L2Discover[9]. Of these three network monitoring tools, only L2Discover has its source code openly available. The company SolarWinds has an open vote on their website on whether or not to include this feature in their monitoring software since 2014[3]. LorientPro and L2Discover use the SNMP for their topology discovery, with STP only being used to discover duplicate and unused links. STPViz uses only STP, thus not needing the networking hardware being SNMP capable.

While commercial tools mostly use active methods for network discovery, several research projects have also attempted to develop purely passive or hybrid tools. One of the first of these tools was developed by Becker et al.[1] in 1995. In more recent years, Blue et al.[2] and Wongsuphasawat et al.[20] published papers on passive network visualization tools in 2008 and 2009. However, we did not find any work on using STP exclusively.

2.2 Spanning Tree Protocol (STP)

Port States

The most important part of STP is its introduction of port states. These port states ensure that the spanning tree does not contain loops. Ports can be in one of three states (see Figure 2.1):

- **Root Port:** the port leading to the root or upwards in the tree. Every non-root bridge has **exactly one** root port.
- **Dedicated Port:** dedicated ports are ports where packets are sent. These are the ports leading "downward" in the tree. The root has only dedicated ports.
- **Blocking Port:** no packets are sent via a blocking Port. This state is used to disable alternate paths to the root or sideways in the tree. In the original paper[15], this state was called backup.

To handle necessary port transitions (e.g. on bridge failure), so called transition states are used. A blocking port for example waits *forwardDelay* seconds in a loop. Every time a superior BPDU is received, the timer is reset. If the timer ever hits zero, the port will transition to dedicated.

In the original STP paper[15], the term Local Area Network (LAN) referred to any connection to a bridge, as well as between bridges. Because the STP information must be propagated to every LAN, there needs to be exactly one dedicated port in a LAN. In a LAN between two non-root bridges, the bridge with the smaller root path cost (see Figure 2.2) will make its connected port dedicated. Should both bridges have the same path cost to the root, the one with the smaller bridge identifier becomes the dedicated bridge (see Section 2.2: STP Packets). This can be seen in the connection between the bottom bridges in Figure 2.1. If no superior STP packets are received on a blocking or root port for a set amount of time (called the forward delay) it transitions to the dedicated state.

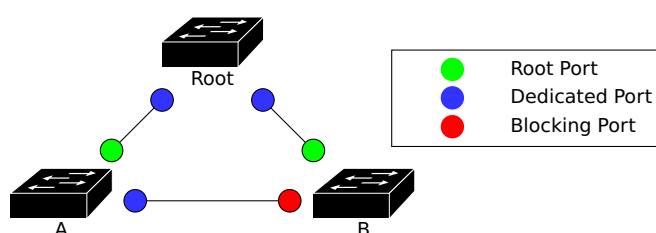


Abbildung 2.1: An example of STP port states

STP Packets

In this section we will explain the individual fields in an STP Bridge Protocol Data Unit (BPDU), as seen in Figure 2.2.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31								
Protocol Identifier																Protocol Version Id								BPDU Type															
Flags						Root Identifier																																	
Root Identifier																																							
Root Identifier						Root Path Cost																																	
Root Path Cost						Bridge Identifier																																	
Bridge Identifier																																							
Bridge Identifier						Port Identifier																Message Age																	
Message Age						Max Age																Hello Time																	
Hello Time						Forward Delay																																	

Abbildung 2.2: An STP BPDU

The fields we used for STPViz are:

- **Flags:** The flags byte is used for the *topology change (TC)* and *topology change acknowledgement (TCA)* flags.
- **Root/Bridge Identifier:** The Identifier consists of three parts and has the same layout for the root and regular bridges:
 - Priority (4 Bits): A value between 0 and 61440 configurable in increments of 4096
 - System ID Extension (12 Bits): Used for keeping the bridge ID unique if multiple VLANs are configured for a bridge
 - Bridge MAC (6 Bytes): The MAC address of the bridge

The conjunction of these three parts is used in comparisons as one large eight byte number.

- **Root Path Cost:** The root path cost is the sum of all port costs (which can be configured or detected automatically) along the current path. The root path cost in packets sent by the root is zero.
- **Message Age:** The message age is the number of bridges that have been passed (in addition to the root) along the current path.
- **Max Age:** If the message age surpasses this maximum, packages are not forwarded any more.
- **Hello Time:** The delay in seconds between sent *Hello* BPDUs. *Hello* BPDUs have the form shown in Figure 2.2.
- **Forward Delay:** After this time, if no superior packet was received, a root, or blocking port will transition to dedicated. Note that if a root port transitions to dedicated state that means the bridge now assumes it is the root.

A BPDU only contains information about the root and the bridge that sent the package. This means that if there is a bridge between these two, we will not know about its ID. We will however still know that it is there, because of the message age. During the buildup of the tree it is possible to obtain information on intermediate nodes. This possibility is explained more in-depth in the section on packet handling (Section 3.2).

Default Parameters

The parameters discussed default to the following values[4]:

- **Priority:** 32768
- **System Id Extension:** 0
- **Max Age:** 20
- **Forward Delay:** 15
- **Hello Time:** 2

Spanning Tree Algorithm

The original paper on STP does not describe an algorithm itself, it merely lists constraints that have to be fulfilled. Algorithm 1 shows the algorithm we use in our own *software-switch* testing tool. We use a reduced number of port states, ignoring the transition states, by utilizing the timestamps of received BPDUs. The algorithm's purpose is to keep loops out of the network, as well as keeping the tree topology consistent.

Data:

received = received packet

current = data of the current bridge

```
if received.rootId < current.rootId then
    /* There is a new root in the network */
    current.rootId = other.rootId
    set receiving port as root-port
    set other ports to dedicated
end
if received.rootPathCost < current.rootPathCost then
    /* This means that the path via the other bridge is shorter, so this
       should be our new root-path */
    set receiving port as root-port
    set other ports to dedicated
else if received.rootPathCost == current.rootPathCost ∧ received.bridgeId <
current.bridgeId then
    /* Both bridges are equidistant from the root, but the other bridge
       has a lower bridge Identifier and should be the dedicated bridge on
       the connection */
    set receiving port as blocking port
end
/* Check if the transitions described in the section on STP Packets are
   necessary */
doPacketTimeOutTransitions()
```

Algorithm 1: Spanning Tree Algorithm (STA)

Extensions

Since the inclusion of STP in the IEEE 802.1D[7] standard, multiple additions have been made to the STP. Rapid STP (RSTP) has been developed to decrease the time needed for tree convergence. Per VLAN STP (PVSTP) has been developed by Cisco and is a proprietary variant which allows for multiple spanning trees to exist, one for each VLAN. The Multiple STP (MSTP) is an extension of RSTP which, like PVSTP, allows for multiple spanning trees to exist in their own VLAN. In addition to that, it also allows for VLAN specific trees to be combined into one large spanning tree.

As an extension to STP the Shortest Path Bridging (SPB)[6] protocol was developed. It allows for multiple paths to be active at the same time, and uses global information to allow

for the calculation of shortest paths.

While the aforementioned extensions to STP would make our task a lot easier (especially SPB), they are not as prevalent as basic STP (which most switches are capable of). For this reason we decided to confine our efforts to basic STP.

2.3 Technologies used

PCAP

PCAP[18] is a shortening of packet capture. It used to be a part of the *tcpdump* tool before it was pulled into it's own library. We used the UNIX version *libpcap* for this thesis. *Libpcap* is a C library and can be directly linked with C and C++ without using a wrapper. It provides functions for opening live network devices and *.pcapng* files. After starting a live capture on a device, *libpcap* will call a callback function for every packet received on the opened interface. The prototype of the callback function looks as follows:

Listing 2.1: Pcap Callback Prototype

```
1 typedef void (*pcap_handler)(u_char *user, const struct pcap_pkthdr *h,  
    const u_char *bytes);
```

Where the parameters are:

- **u_char *user** is used to pass user defined parameters.
- **pcap_pkthdr *h** contains useful information about the packet, like source and destination addresses, as well as size.
- **u_char *bytes** contains the actual (binary) data of the packet.

The **typedef void (*pcap_handler)** part of the prototype just means that the function returns a **void ***.

Usage of this function will be covered more in-depth in the chapter on STPViz itself (Chapter 3).

JSON

JSON stands for Java Script Object Notation and was used for two reasons:

1. It keeps the network communication independent from any programming languages used.
2. Utility libraries for JSON are easily available for most programming and scripting languages, saving us the work of inventing and implementing our own notation.

JSON has a simple notation for declaring objects and arrays, as well as primitive data types. An example is shown in Listing 2.2.

Listing 2.2: JSON Example

```
1 { "users":[
2   {
3     "firstName":"Ray",
4     "lastName":"Villalobos",
5     "joined": {
6       "month":"January",
7       "day":12,
8       "year":2012
9     }
10  },
11  {
12    "firstName":"John",
13    "lastName":"Doe",
14    "joined":{
15      "month":"April",
16      "day":28,
17      "year":2010
18    }
19  }
20 ]}
```

The JSON library used in this thesis is *jsoncpp*[13].

TikZ

We chose TikZ ist kein Zeichenprogramm (TikZ)[17] as the file format for our output. Using TikZ allows us to generate .tex files rather than images. These are easier to generate, as well as to modify after creation. It is also very easy to integrate in L^AT_EX papers. While TikZ is very powerful and therefore complex, the parts we use are fairly simple. Listing 2.3 shows the TikZ code for Figure 1.1b.

Listing 2.3: A TikZ Example

```
1 \begin{tikzpicture}
2   \node (root) at (4,8) {\switch{0.8}{A}};
3   \node (B) at (2,6) {\switch{0.8}{B}};
4   \node (C) at (6,6) {\switch{0.8}{C}};
5   \node (D) at (4,4) {\switch{0.8}{D}};
6
7   \draw
8     (root) edge (B)
9     (root) edge (C);
10
11  \draw[green, thick, ->]
12    (B) edge (D)
13    (C) edge (D);
```

```
14
15     \draw[red, thick, <->]
16     (B) edge (C);
17 \end{tikzpicture}
```

The `\node` keyword draws a node with the id given in parenthesis. It is centered around the coordinates given with the *at* keyword and has the text written in braces. The (0,0) coordinate is in the top left corner of an image. For our graphics we used a self defined `\switch` macro, but regular text or \LaTeX commands work as well. Edges are drawn using the `\draw` command. This command can be used with raw coordinates or identifiers. Options for the edges can be passed in brackets. Note that *edge* can be substituted with `- -`. All TikZ commands are ended with a semicolon. TikZ graphics are automatically cropped on rendering.

Kapitel 3

STP Tree Generator

We split the application in three parts:

- **Client:** collects STP information and sends it to the server. The client also handles piecing together the packets into paths.
- **Server:** saves the data from the clients and combines them into one tree.
- **Parser:** contacts the server to receive the tree and converts it into output format.

The intended form of usage is to have multiple clients in the network connecting to one server. We combine information received from multiple clients to reach a better understanding of the network.

STP uses only local data, which means that bridges have no knowledge of the network, except for their own port states. Unfortunately, this makes it hard to find connections between bridges. The only way to obtain this information is by capturing packets during the tree build up.

When the client witnesses the message age increasing, it assumes the new root to be prepended to the previous one. This assumption is uncertain, as this connection cannot be guaranteed, but the risk of making mistakes can be reduced. Details on this are discussed in the section on uncertain assumptions (Section 3.4)

Figure 3.1 shows a rudimentary example of information gained during tree build up. As bridge *B* has not yet received an STP packet from bridge *A* yet, it thinks itself root, so it sends STP packets with its ID as bridge and root identifier (Figure 3.1a). The client therefore only knows about bridge *B*. When bridge *B* receives an STP packet from *A*, it updates its local information and broadcasts it, as shown in Figure 3.1b. This leads to the client adding bridge *A* to its vector of known bridges and increments the saved message age for all other bridges. Both bridges are contained in an STP packet from *B* as root and bridge respectively, but the process stays the same for larger topologies.

We only assume network growth for the cases where the message age increases at the same time that the root changes. For other cases (message age decreasing or staying constant) we did not find ways to gain knowledge about the network. Because we cannot guarantee that the topology change does not affect the path we assume, STPViz clients currently clear their data unless the message age increases. With the limited time and networking hardware at our disposal, we were not able to find a way to reduce the amount of information that has

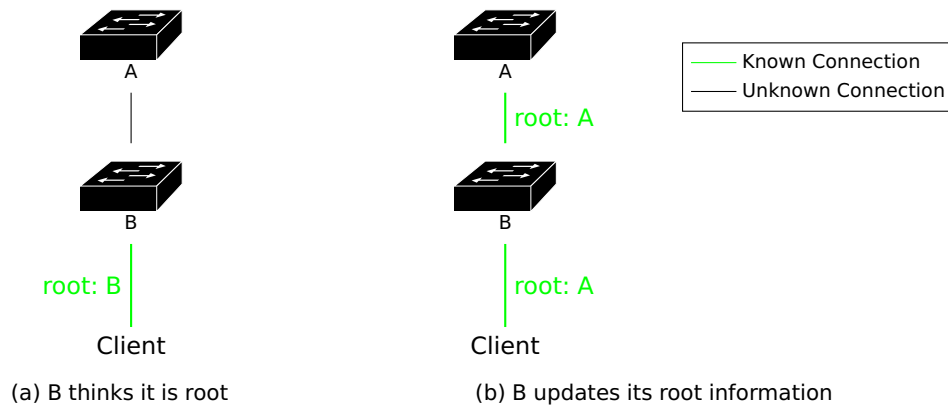


Abbildung 3.1: Information gained on STP build up

to be discarded. We believe that by increasing the communication between clients, STPViz would be able to only discard information that is known to be false.

Figure 3.2 shows an example case where the client has to reset its data. In Figure 3.2a the client knows about the full path to the current root, bridge A. When bridge C is connected (before Figure 3.2b) and becomes root, a TC flag is sent along with a different message age. Due to the message age decreasing when the root changes, STPViz discards all data and adds the two bridges contained in the STP packet again. This leaves it with knowledge of bridges D and C. A detailed explanation of how STPViz handles incoming STP packets can be found in the section on packet handling (Section 3.2).

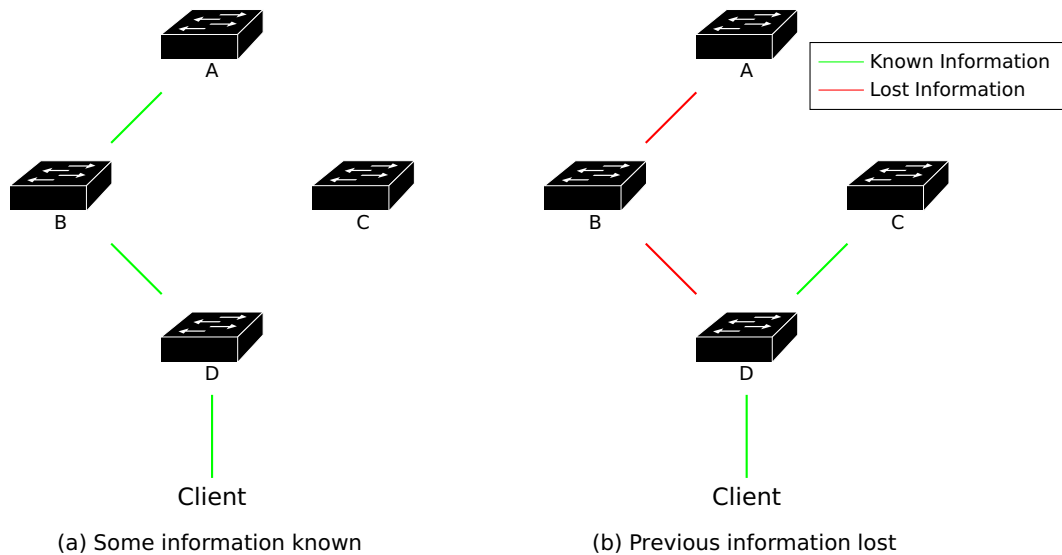


Abbildung 3.2: Information lost during tree build up

3.1 Class Structure

The classes we created made saving the STP data easier and reduced the effort needed to generate the *JSON* and *TikZ* output. We created classes to represent MAC addresses, bridges, and complete spanning trees. All these classes have conversion functions from and to *JSON* format, as well as a function to generate *TikZ* output. Altogether the following classes were created:

- **Mac**: A container class for MAC addresses. It is used to store the address in an easier to use format (as a C++ `std::string`).
- **Bridge**: Stores a **Mac** object in conjunction with the priority and message age.
- **SpanningTree**: This class represents an entire tree. It has functions for creating the *TikZ* export as well as combining and manipulating subtrees. The **SpanningTree** class is a recursive data structure, storing the root as a *Bridge* object and its children as a vector of **SpanningTree** objects. As a lot of its functions are needed in this chapter, the header file is shown in detail in Listing 3.1.
- **Sniffer**: Does the actual packet sniffing.
- **Client**: Handles the client side communication.
- **Server**: Handles the server side communication, as well as combining the trees and removing incorrect information.

Listing 3.1: The **SpanningTree** header file

```
1 #pragma once
2 #include <vector>
3 #include <cassert>
4 #include <iostream>
5 #include <string>
6 #include <utility>
7
8 #include "Bridge.hpp"
9 #include "../inc/json/json.h"
10 #include "../inc/json/json-forwards.h"
11
12 class SpanningTree{
13     private:
14         Bridge root;
15         std::vector<SpanningTree> children;
16         std::pair<std::string, int> toTikzHelper(double lowerX, double
            upperX, int y, int yStep, int oldMessageAge, int index) const;
17
18     public:
19         SpanningTree();
20         SpanningTree(const Bridge& r);
```

```

21     SpanningTree(const Bridge& r, std::vector<SpanningTree> c);
22     SpanningTree(const SpanningTree& other);
23     ~SpanningTree();
24     void addChild(const SpanningTree& child);
25     int addSubTree(const SpanningTree& other);
26     int containsRoot(const SpanningTree& tree) const;
27     Json::Value toJson() const;
28     std::string toTikz(double lowerX, double upperX, int Y, int yStep,
29         int oldMessageAge, int index) const;
29     int maxWidth() const;
30     int widthAtLevel(int levelsRemaining) const;
31     int size() const;
32     Bridge getRoot() const;
33
34     static SpanningTree fromJson(const Json::Value buildFrom);
35
36     friend SpanningTree operator+(const SpanningTree& lhs, const
37         SpanningTree& rhs);
37     friend int operator==(const SpanningTree& lhs, const SpanningTree&
38         rhs);
38     friend int operator!=(const SpanningTree& lhs, const SpanningTree&
39         rhs);
39     SpanningTree operator=(const SpanningTree& other);
40 };

```

The STP data is saved in the **Sniffer** as a vector. This is easier than storing them in a fixed size array, and already combining them into a **SpanningTree** object would keep the server from performing the steps described in Section 3.4.

3.2 Packet Handling

Packets are handled by the **Sniffer** class. While the function mostly skips unused fields in the STP packet, here we will take a close look at the more important parts. In order to check if the packet is actually an STP packet, we use the Ethernet destination address. Listing 3.2 shows how that is done. The *bytes* variable is provided by the required *pcap* callback prototype (see Section 2.3: PCAP). This way is computationally more expensive than saving the destination in binary format and comparing memory. It is however also more readable and a lot easier to change should the need arise. The address we compare the destination to is the broadcast target for STP.

Listing 3.2: Filtering for STP packets

```

1     struct ethhdr *ethh = (struct ethhdr*) bytes;
2
3     char buffer[17];
4     sprintf(buffer, "%.2X:%.2X:%.2X:%.2X:%.2X:%.2X", ethh->h_dest[0], ethh
5         ->h_dest[1], ethh->h_dest[2], ethh->h_dest[3], ethh->h_dest[4],

```

```

        ethh->h_dest[5]);
5     if(std::string(buffer, 17) != "01:80:C2:00:00:00")
6         return;

```

Pcap provides us with a pointer to the binary packet data. We can go through the packet by incrementing this pointer by set amounts (after skipping to the STP data). This repeats for the bridge identifier, as well as other fields. Listing 3.3 shows the beginning of the payload handling.

Listing 3.3: Going Through the Payload

```

1     //handle payload
2     const u_char *payload = bytes+sizeof(struct ethhdr);
3     int psize = header->len - sizeof(struct ethhdr);
4     //payload starts at logical link control level
5     //llc has 3 bytes
6     payload+=3;
7     psize-=3;
8     //payload is now at the start of the stp payload
9     //protocol identifier stp (2 bytes)
10    payload+=2;
11    psize-=2;
12
13    //one byte for version and bpdu type
14    payload+=2;
15    psize-=2;
16
17    //tc flag
18    bool tc = *(payload++);
19    psize--;
20
21    //root identifier
22    //bridge priority is the next 2 bytes
23    //actual bridge priority is 4 bits,
24    //then comes id extension with 12 bits
25
26    //generate filters
27    unsigned short extFilter = -1;
28    extFilter = extFilter>>4;
29    unsigned short prioFilter = extFilter<<12;
30
31    unsigned short rPriority, rExtension;
32    rPriority = htons(*((short*)payload));
33    rExtension = rPriority & extFilter;
34    rPriority = rPriority & prioFilter;
35    payload+=2;
36    psize-=2;
37    //next 6 bytes is the root bridge id

```

```

38     Mac rootMac(payload);
39     payload+=6;
40     psize-=6;
41
42     //root path cost (4 bytes)
43     payload+=4;
44     psize-=4;
45
46     //bridge identifier
47     //bridge priority is the next 2 bytes
48     unsigned short bPriority, bExtension;
49     bPriority = htons(*((short*)payload));
50     bExtension = bPriority & extFilter;
51     bPriority = bPriority & prioFilter;
52     payload+=2;
53     psize-=2;
54     //next 6 bytes is the bridge mac address
55     Mac bridgeMac(payload);
56     payload+=6;
57     psize-=6;
58
59     //next 2 bytes are port identifier
60     payload+=2;
61     psize-=2;

```

After the data is extracted from the packet, it is constructed into our custom classes. We then check if the two bridges (root and first hop) were previously known, as seen in Listing 3.4. Concerning the root, knowing whether or not it was previously known is enough. For the first hop, we also require knowledge about the old message age.

Listing 3.4: Checking for Previously Known Information

```

1     //next 2 bytes is message age
2     unsigned short messageAge;
3     messageAge = *((short*)payload);
4     payload+=2;
5     psize-=2;
6
7     //generate Bridge objects
8     Bridge root(rootMac, rPriority, rExtension, 0);
9     Bridge firstHop(bridgeMac, bPriority, bExtension, messageAge);
10
11    //check if the two nodes are contained
12    int rootContained = 0, oldHopMa = -1;
13    for(Bridge b : bridges){
14        if(b == root)
15            rootContained = 1;
16        if(b == firstHop)

```

```

17         oldHopMa = b.getMessageAge();
18     }

```

Listing 3.5 shows how we update the bridge information in the sniffer. The *clearAndAdd* function is just a shorthand to clear the bridge vector before adding the two bridges to it.

Listing 3.5: Bridge Data Update

```

1  //handle network changes
2  if(rootContained){
3      if(oldHopMa >= 0){
4          //both contained
5          if(oldHopMa != firstHop.getMessageAge()){
6              //changes occurred
7              clearAndAdd(firstHop, root);
8          }else{
9              //everything is as it was
10             //if something changed upstream a tc flag will be sent
11             //->reset
12             if(tc && !hadTC)
13                 clearAndAdd(firstHop, root);
14         }
15     }else{
16         //first hop not contained
17         //this means that the node was plugged in on a different
18         bridge
19         //(most likely)
20         //we don't know how long we were disconnected, so reregister
21         the client
22         if(!noConnect)
23             client->regServer();
24         clearAndAdd(firstHop, root);
25     }
26 }else{
27     //root not contained
28     if(oldHopMa >= 0){
29         //if the first hop was contained this means there were some
30         changes upstream
31         //the root moving away means we assume network growth
32         if(oldHopMa < firstHop.getMessageAge()){
33             std::cout << "root moved away\n";
34             int maDiff = firstHop.getMessageAge() - oldHopMa;
35             //this means we have to increase every message age and add
36             the new root
37             for(Bridge &b : bridges)
38                 b.setMessageAge(b.getMessageAge() + maDiff);
39             bridges.push_back(root);
40         }else{

```

```

37         clearAndAdd(firstHop, root);
38     }
39 }else{
40     //entirely new setup
41     clearAndAdd(firstHop, root);
42 }
43 }

```

3.3 Communication

Our distributed architecture is hierarchical, as we have a distinct *client-server* separation. We also use a *push* style communication, meaning the clients will push their messages to the server, which simply waits for incoming connections. The alternative would be to have the server poll the registered clients for their data. This would have doubled the communication effort needed and made the handling for disconnected clients more complicated.

As discussed in the background section (Section 2.3) we are using JSON for the client server communication. To inform all involved components about the purpose of a JSON message, we use a *messagetype* field. The possible values for *messagetypes* are as follows:

- **Register:** before they can send data to the server, the clients need to register to receive an identifier. These unique identifiers are used to keep the data from different clients separated.
- **Push:** clients send *push* messages when they transmit their data to the server. These messages contain lists of bridges. They are transmitted the same way that they are stored in the sniffer, without modifications.
- **Report:** the parser sends a message to the server containing this *messagetype* and nothing else. The server then combines the client data and transmits it to the parser.

Bridge data is transmitted from the client to the server in standard JSON array notation. When the data is transmitted from the server to the parser, it is transmitted as a full tree. We added conversion functions to and from JSON to all our custom data classes, to keep this transmission simple. Listing 3.6 shows an example transmission from client to server, and Listing 3.7 shows a transmission from server to parser. The JSON tree shown in Listing 3.7 is one of the trees transmitted during our test runs. It shows how STPViz can use changing packets to gather information about bridge connections.

Listing 3.6: Client-Server Transmission

```

1 {
2   "bridges" :
3   [
4     {
5       "mac" :
6       {
7         "address" : "F4:F2:6D:7D:BF:BD"

```

```

8      },
9      "messageAge" : 0,
10     "priority" : 144
11   }
12 ],
13 "messagetype" : "push"
14 }

```

Listing 3.7: Server-Parser Transmission

```

1 {
2   "children" :
3   [
4     {
5       "children" :
6       [
7         {
8           "children" : null,
9           "root" :
10          {
11            "mac" :
12            {
13              "address" : "F4:F2:6D:7D:BF:BD"
14            },
15            "messageAge" : 2,
16            "priority" : 240
17          }
18        }
19      ],
20      "root" :
21      {
22        "mac" :
23        {
24          "address" : "00:09:6B:93:28:2E"
25        },
26        "messageAge" : 1,
27        "priority" : 384
28      }
29    },
30    {
31      "children" : null,
32      "root" :
33      {
34        "mac" :
35        {
36          "address" : "50:7B:9D:0C:9F:CD"

```



```

37         },
38         "messageAge" : 1,
39         "priority" : 384
40     }
41 }
42 ],
43 "root" :
44 {
45     "mac" :
46     {
47         "address" : "BC:AE:C5:EB:7D:B6"
48     },
49     "messageAge" : 0,
50     "priority" : 128
51 }
52 },
53 {
54     "children" : null,
55     "root" :
56     {
57         "mac" :
58         {
59             "address" : "BC:AE:C5:EB:7D:B6"
60         },
61         "messageAge" : 0,
62         "priority" : 128
63     }
64 }

```

3.4 Uncertain Assumptions

As previously stated, the assumptions we make about bridge connections are not necessarily true. They can be wrong, for certain cases, one of which is shown in Figure 3.3. The client knows about the path to (at the time) root *B* in Figure 3.3a. When the global root information gets updated, the message age will increase causing the client to prepend *A* to its known path, even though there is no connection between *A* and *B*. This leads the client to thinking the path looks like in Figure 3.3b.

We did not find an alternate way to gather information about connections in the network, so we tried to reduce errors created by our method. To this end we combine information gathered by multiple clients to identify and remove false information. If a client makes a false assumption about a bridge, the message age for that bridge will be lower than its actual message age. An example can be seen in Figure 3.4. It uses the same topology as Figure 3.3. The assumed (nonexistent) connection between bridges *A* and *B* is shown in red. Figure 3.4a shows the message ages assumed by STPViz. With the assumed connection the message age of bridge *B* would be one. However, as Figure 3.4b shows, the actual message age is three.

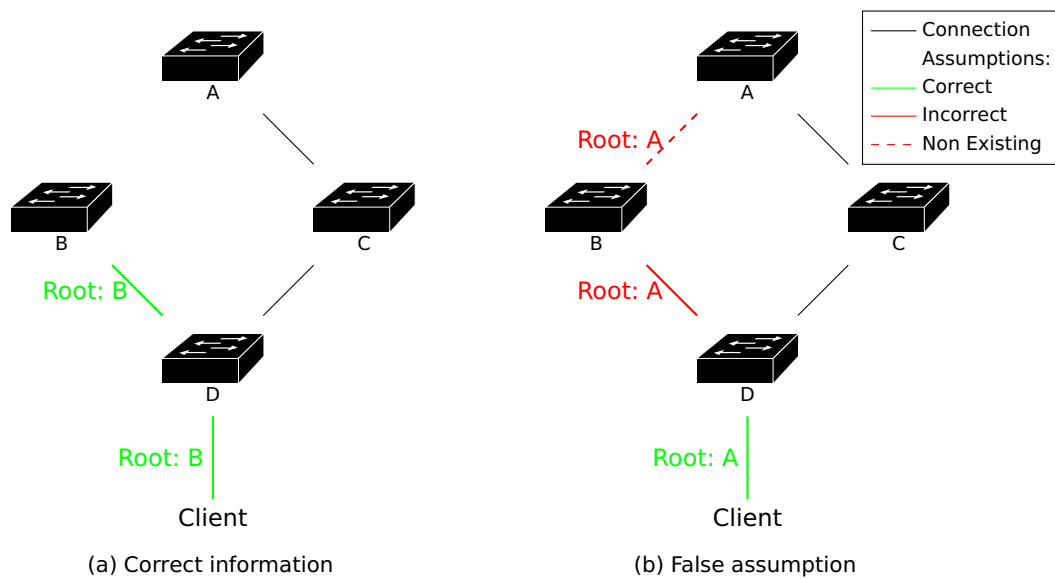


Abbildung 3.3: An example case of false assumptions about network structure

If a client connected to bridge *B* sends its actual message age to the server (or any message age larger than one), the server will remove *B* from the path sent by the client.

By checking all bridges for inconsistencies we can remove any false assumptions from the data we use for building the visualization. The code for doing this is shown in Listing 3.8.

Listing 3.8: Removing Incorrect Bridge Data

```

1  //remove invalid entries before creating SpanningTree
2  //check every received vector against every other
3  for(auto mapIt1 = clientData.begin(); mapIt1 != clientData.end(); ++
    mapIt1){
4      std::vector<Bridge> vec1 = mapIt1->second;
5      for(auto mapIt2 = clientData.begin(); mapIt2 != clientData.end();
        ++mapIt2){
6          if(mapIt1 == mapIt2)
7              continue;
8          std::vector<Bridge> vec2 = mapIt2->second;
9          for(auto vecIt1 = vec1.begin(); vecIt1 != vec1.end(); ++vecIt1
            ){
10             for(auto vecIt2 = vec2.begin(); vecIt2 != vec2.end(); ++
                vecIt2){
11                 //check every bridge against every other
12                 //if one has a smaller message age, it is incorrect
                    and needs to be removed
13                 //this might overlook some errors if the number of
                    clients is insufficient
14                 if(vecIt1->getMessageAge() < vecIt2->getMessageAge()){

```

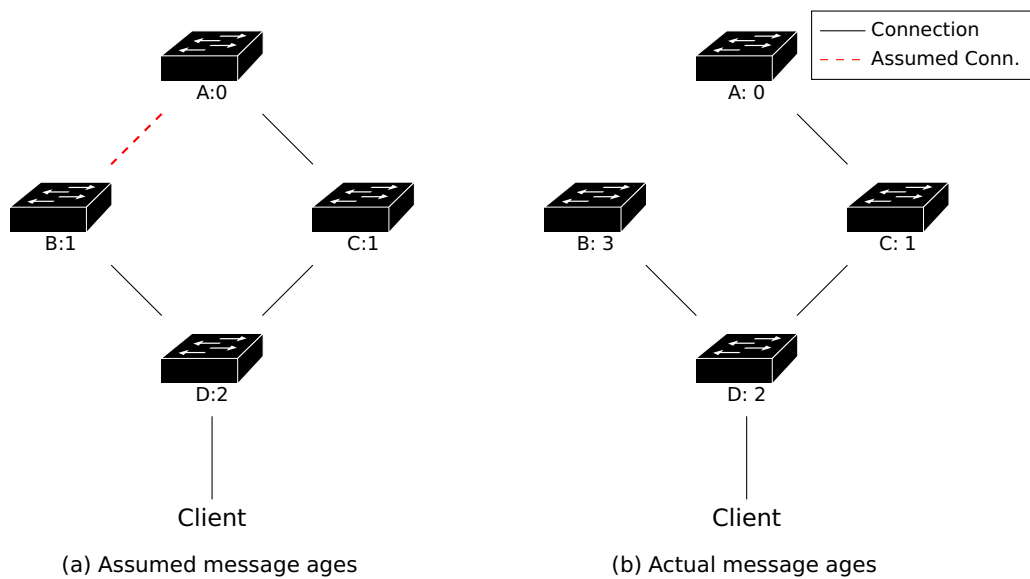


Abbildung 3.4: Assumed and actual message ages

```

15         //remove bridge
16         vecIt1=vec1.erase(vecIt1);
17         //decrement vecIt1 to have it at the correct
           position in the next loop iteration
18         --vecIt1;
19         //continue with outer loop (vec1 iteration)
20         break;
21     }else if(vecIt2->getMessageAge() < vecIt1->
           getMessageAge()){
22         vecIt2=vec2.erase(vecIt2);
23         --vecIt2;
24         //no break needed -> next step is inner loop (vec2
           iteration)
25     }
26 }
27 }
28 }
29 }
```

The bridges are removed from the vector they each are in. This way any connection information that is not proven wrong is retained, which means that if previous assumptions were correct, they will not be lost.

3.5 Combining the data

After false information is removed, the bridge data is combined into one `SpanningTree` object. Creating large `SpanningTree` objects is done in multiple steps. First, we combine the bridge vectors to individual trees. These trees each represent a path from a client to the root. After the individual trees are obtained, we combine the ones with the same root. The code for this can be seen in Listing 3.9.

Listing 3.9: Combining the Bridge Data

```
1  //now that the incorrect entries are removed, we assume that
   //everything else is correct
2  //WHICH MAY STILL BE WRONG
3  //create SpanningTrees like in Sniffer
4  std::vector<SpanningTree> indivTrees = std::vector<SpanningTree>();
5  for(auto mapIt = clientData.begin(); mapIt != clientData.end(); ++
   mapIt){
6      auto vec = mapIt->second;
7      std::sort(vec.begin(), vec.end(), [](Bridge a, Bridge b) {return a
   .getMessageAge() < b.getMessageAge();});
8      if(vec.size() > 0)
9          indivTrees.push_back(treeHelper(vec.begin(), vec.end()));
10 }
11
12 //now combine these trees
13 std::vector<SpanningTree> ret;
14 for(auto indivIt = indivTrees.begin(); indivIt != indivTrees.end(); ++
   indivIt){
15     bool contained = false;
16     for(auto treeIt = ret.begin(); treeIt != ret.end(); ++treeIt){
17         if(treeIt->containsRoot(*indivIt) && *treeIt != *indivIt){
18             *treeIt = *treeIt + *indivIt;
19             contained = true;
20             break;
21         }
22     }
23     if(!contained)
24         ret.push_back(*indivIt);
25 }
```

Combining the individual trees into one is done using the operator we overloaded. This operator executes a recursive tree combine, which is shown in Figure 3.10.

Listing 3.10: Recursive Tree Combination Algorithm

```
1 SpanningTree operator+(const SpanningTree& lhs, const SpanningTree& rhs){
2     if(lhs == rhs){
3         return SpanningTree(lhs);
```

```

4     }
5
6     assert(lhs.root == rhs.root);
7     SpanningTree ret = SpanningTree(lhs.root);
8
9     //add all similar children
10    for(SpanningTree leftChild : lhs.children){
11        for(SpanningTree rightChild : rhs.children){
12            //if they have the same root, add the combination
13            if(leftChild.root == rightChild.root){
14                //should we make an incorrect assumption, set the message
                //age to the higher one.
15                //this should not be needed, as this should be filtered in
                //the server.
16                SpanningTree newTree = leftChild + rightChild;
17                newTree.root.setMessageAge(std::max(leftChild.root.
                    getMessageAge(), rightChild.root.getMessageAge()));
18                ret.addChild(newTree);
19                break;
20            }
21
22            //if rightChild is contained in left (additions "downstream"),
                //add it there
23            if(leftChild.containsRoot(rightChild)){
24                SpanningTree combined = SpanningTree(leftChild);
25                leftChild.addSubTree(rightChild);
26                ret.addChild(combined);
27            }
28
29            //vice versa
30            if(rightChild.containsRoot(leftChild)){
31                SpanningTree combined = SpanningTree(rightChild);
32                leftChild.addSubTree(leftChild);
33                ret.addChild(combined);
34            }
35        }
36    }
37
38    //add all remaining children of lhs
39    for(SpanningTree leftChild : lhs.children){
40        int contained = 0;
41        for(SpanningTree added : ret.children){
42            if(added.root == leftChild.root){
43                contained = 1;
44                break;
45            }
46        }

```

```

47         if(!contained){
48             ret.addChild(leftChild);
49         }
50     }
51     //add all remaining children of rhs
52     for(SpanningTree rightChild : rhs.children){
53         int contained = 0;
54         for(SpanningTree added : ret.children){
55             if(added.root == rightChild.root){
56                 contained = 1;
57                 break;
58             }
59         }
60         if(!contained){
61             ret.addChild(rightChild);
62         }
63     }
64
65     return ret;
66 }

```

Our Tree combination algorithm uses a helper function called *addSubTree*, which is shown in Listing 3.11.

Listing 3.11: Adding subtrees in the place where they belong

```

1  int SpanningTree::addSubTree(const SpanningTree& other)
2  {
3      if(!containsRoot(other))
4          return 0;
5
6      if(root == other.root){
7          //combine all similar children
8          for(SpanningTree leftChild : children){
9              for(SpanningTree rightChild : other.children){
10                 if(leftChild.root == rightChild.root){
11                     leftChild.addSubTree(rightChild);
12                     break;
13                 }
14             }
15         }
16
17         //add all missing right children to the tree
18         for(SpanningTree rightChild : other.children){
19             int contained = 0;
20             for(SpanningTree leftChild : children){
21                 if(leftChild.root == rightChild.root){
22                     contained = 1;

```

```

23             break;
24         }
25     }
26     if(!contained)
27         addChild(other);
28 }
29 }
30
31 return 1;
32 }

```

The comparison operator we overloaded for our **SpanningTree** class recursively compares the roots of the subtrees for equality (which in turn just compares all members). It is shown in Listing 3.12.

Listing 3.12: Comparing SpanningTree objects for equality

```

1 int operator==(const SpanningTree& lhs, const SpanningTree& rhs){
2     if(lhs.root != rhs.root)
3         return 0;
4
5     //check if both have the same amount of children (if not, they are not
6     //the same)
7     if(lhs.children.size() != rhs.children.size())
8         return 0;
9
10    //check if all children of lhs are contained in the children of rhs
11    //if this holds, all children of rhs are going to be contained in lhs
12    //(because they have the same amount of children)
13    for(const SpanningTree& childL : lhs.children){
14        int contained = 0;
15        for(const SpanningTree& childR : rhs.children){
16            if(childL == childR){
17                contained = 1;
18                break;
19            }
20        }
21        if(!contained)
22            return 0;
23    }
24    return 1;
25 }

```

3.6 Visualization

STPViz generates *.tex* files containing *TikZ* graphs. Nodes in these graphs contain the bridge data. Edges show the (assumed) connections in the topology. Node information is displayed in the format *priority, system id extension - MAC address, message age*. For cases where STPViz has no knowledge of the path to the root, empty nodes are used. These can be seen as gaps in the final output. Figure 3.5 shows an example.

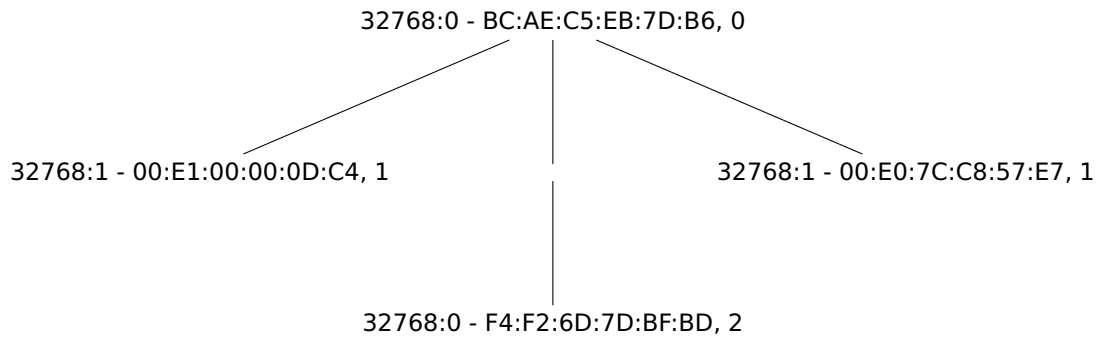


Abbildung 3.5: An example for STPViz output

The output is generated using a recursive function in the **SpanningTree** class. Our code for this is rather confusing when first read, so we decided to split it into multiple listings, allowing for clearer explanations. No code is skipped between the listings, so combining Listings 3.14-3.17 would yield the original function. To make it easier to follow we also did not modify any indentations. Viewing the code together with the raw \LaTeX output makes it much clearer, so Listing 3.13 shows the code for Figure 3.5

Listing 3.13: *.tex* code for Figure 3.5

```
1 \begin{tikzpicture}[]
2 \node (0) at (7.000000,20) {32768:0 - BC:AE:C5:EB:7D:B6, 0};
3 \node (1) at (3.500000,18) {32768:1 - 00:E0:7C:C8:57:E7, 1};
4 \node (2) at (3.500000,16) {61440:0 - F4:F2:6D:7D:BF:BD, 2};
5 \draw
6 (1) -- (2);
7 \node (3) at (10.500000,18) {32768:1 - 00:E1:00:00:0D:C4, 1};
8 \draw
9 (0) -- (1)
10 (0) -- (3);
11 \end{tikzpicture}
```

For an explanation on TikZ keywords please refer to the section on TikZ (Section 2.3).

Listing 3.14: TikZ conversion function, part 1

```
1 std::string SpanningTree::toTikz(double lowerX, double upperX, int y, int
    yStep, int oldMessageAge, int index) const {
```



```

2     return toTikzHelper(lowerX, upperX, y, yStep, oldMessageAge, index).
        first;
3 }
4
5 std::pair<std::string, int> SpanningTree::toTikzHelper(double lowerX,
    double upperX, int y, int yStep, int oldMessageAge, int index) const {
6     //this helper is needed because the previously used child index
        calculation was faulty if the algorithm was adding empty nodes
7     std::string retString;
8     int retInt = index;
9
10    double dist = upperX - lowerX;
11    double x = lowerX + dist/2;

```

The *lowerX*, *upperX*, *y* and *yStep* parameters are used to calculate the position for the current root, as well as the subtrees. *OldMessageAge* is used to discern whether or not STPViz needs to draw empty nodes. As we use incrementing IDs to draw connections between the nodes (see Listing 3.13), we need to *index* to pass the current index. This function is executed recursively until the current **SpanningTree** has no children, so the first step is always to calculate the position of the current root (see Listing 3.1). The previous child index calculation will be discussed in the explanation for Listing 3.17.

Listing 3.15: TikZ conversion function, part 2

```

1     if(root.getMessageAge() > oldMessageAge+1){
2         //if there is an unknown node between the previous node and this
            one create an empty node
3         retString += "\\node (" + std::to_string(index) + ") at (" + std:::
            to_string(x) + "," + std::to_string(y) + ") {};\n";
4         std::pair<std::string, int> subRet = toTikzHelper(lowerX, upperX,
            y-yStep, yStep, oldMessageAge+1, index+1);
5         retString += subRet.first;
6         retInt = subRet.second;
7         retString += "\\draw (" + std::to_string(index) + ") -- (" + std:::
            to_string(index+1) + ");\n";

```

If there is an unknown bridge between the current root and the previous one, we draw an empty node to visualize this. After the empty node is drawn, the same function will be called with increased *y* coordinate, because the current root has not been handled yet. We also know that the ID will be the current index plus 1. With these informations we can write the command for drawing the edge beforehand. Note that we include newlines to keep the output readable. As the returned string is a C++ **std::string** we use the `\n` escape sequence.

Listing 3.16: TikZ conversion function, part 3

```

1     }else{
2         //draw this node

```

```

3      retString += "\\node (" + std::to_string(index) + ") at (" + std::
      to_string(x) + "," + std::to_string(y) + ") {" + root.toTikz()
      +"};\n";
4
5      //draw children
6      //we need the return values of the children later, for the index
      calculation
7      double xPerWidth = dist / maxWidth();
8      std::vector<std::pair<std::string, int>> childRets;
9      for(unsigned int i=0; i<children.size(); i++){
10         upperX = lowerX + xPerWidth * children[i].maxWidth();
11         if(i==0)
12             childRets.push_back(children[i].toTikzHelper(lowerX,
13                 upperX, y-yStep, yStep, root.getMessageAge(), index+1)
14             );
15         else
16             childRets.push_back(children[i].toTikzHelper(lowerX,
17                 upperX, y-yStep, yStep, root.getMessageAge(),
18                 childRets[i-1].second+1));
19
20         lowerX = upperX;
21     }
22
23     //add the strings in an additional loop, to have nicer formatting
24     of the output
25     for(auto ret : childRets)
26         retString += ret.first;

```

Here we draw the node for the current root. To calculate the amounts of horizontal space the subtrees can use, we use their respective maximum width. This width is calculated using the functions shown in Listing 3.18. For subtree spacing we simply use the relation between the width of the full tree and the width the individual children. In Listing 3.16 we see the reason why we have to return a pair of **string** and **int**. As it would be rather expensive to check for the number of empty nodes that need to be drawn beforehand (this would double the required computational effort), we return the last used index as well. This eliminates ID conflicts. To keep drawing the nodes and edges in blocks, we use multiple loops for drawing them. This leads to edges being drawn in layers", as edges from a node to all its children are drawn at once.

Listing 3.17: Tikz conversion function, part 4

```

1      //draw lines to children
2      if(childRets.size() >0){
3          retString += "\\draw ";
4          for(unsigned int i=0; i<childRets.size(); i++){
5              if(i==0)
6                  retString += "\n(" + std::to_string(index) + ") -- ("
                  + std::to_string(index+1) + ")";

```

```

7         else
8             retString += "\n(" + std::to_string(index) + ") -- ("
                + std::to_string(childRets[i-1].second+1) + ")";
9
10        }
11        retString += ";\n";
12        //set return index to largest children index
13        retInt = childRets[childRets.size()-1].second;
14    }
15 }
16
17 return std::pair<std::string, int>(retString, retInt);
18 }

```

Listing 3.17 shows how we draw edges between nodes and their children. We use the same method of accessing the indices as in Listing 3.16. Before returning we set the returned *int* half of the pair to the largest index used by a recursive call.

Listing 3.18: Calculating the maximum width of a tree

```

1 int SpanningTree::maxWidth() const{
2     int max = 1;
3     for(int level = 0, currentWidth; (currentWidth = widthAtLevel(level))
        > 0; level++)
4         max = max > currentWidth ? max : currentWidth;
5
6     return max;
7 }
8
9 int SpanningTree::widthAtLevel(int levelsRemaining) const{
10     if(levelsRemaining <= 0)
11         return 1;
12     if(levelsRemaining == 1)
13         return children.size();
14
15     int ret = 0;
16     for(SpanningTree s : children)
17         ret += s.widthAtLevel(levelsRemaining-1);
18     return ret;
19 }

```

3.7 Installation & Usage

STPViz was written in C++ and requires a compiler capable of C++11. Usage of the included Makefile is strongly advised. We included the version of *jsoncpp* we used, so it does not need to be downloaded. The client requires an installation of the *pcap* development library to

compile, and the regular library to run. Available build targets are **client**, **server**, **parser** for single components and **all** for all components. An **install** target for global installation is not provided.

Client Launch Parameters

-if inputFile is used to specify the name of an input file which is written in *.pcapng* format. If specified, packets will not be captured live, but instead taken from the input file. For easier usage with this parameter, launching the server with a higher timeout duration is recommended. The server will remove data from this client, as it is not programmed to resend data taken from input files.

-of outputFile is used to specify a different output file. Output files are used for logging data. Logging directly to **stdout** is currently not possible.

-no-connect / -nc will stop the client from connecting to a server. This was used for debugging the client, but remains in the current version.

-h hostname will tell the client which hostname to connect to. IPv4 addresses in dot notation, as well as hostnames (including URLs) are accepted. For more information on accepted formats please refer to the manual page for **gethostbyname**.

-p port specifies the port on which to connect.

-dn deviceName tells the client the device name of the interface to use. The names can be obtained using commands such as **ifconfig** or **ip addr show**.

Server Launch Parameters

-p port specifies the port to listen on.

-of outputFile specifies the filename for the output file, which is used for logging incoming data, as well as the current state.

-np disables the creation of a *.pid* file containing the process id. In cases the server is meant to be launched automatically in the background this file provides an easy means of accessing the process id.

-t timeout tells the server how large the timeout time should be before removing a client's data.

Parser Launch Parameters

-p port specifies the port to connect on.

-h *hostname* specifies the hostname to connect to.

-pw *pictureWidth* specifies the width of the resulting TikZ picture. Note that if a width of 20cm is specified but only 10cm are used, the picture will be snipped automatically

-ph *pictureHeight* specifies the height of the resulting TikZ picture. The same snipping occurs as for the width.

-s *yStep* defines the vertical distance between layers.

Default Parameters

Client parameters default to the following values:

- **-if:** defaults to nothing, live capture is used
- **-of:** client.log in the current working directory
- **-nc:** not set
- **-h:** *localhost*
- **-p:** 80
- **-dn:** first named device

Server parameters default as follows:

- **-p:** 80
- **-of:** server.log in the current working directory
- **-np:** not set
- **-t:** 10

Parser parameters default to the following values:

- **-p:** 80
- **-h:** *localhost*
- **-s:** 2
- **-ph:** 20
- **-pw:** 14

Configuration Files

The binaries all create configuration files with their current configuration if none exist. On subsequent launches they will take their configuration from these files. Command line parameters take precedence over configuration files, which in turn take precedence over default parameters.

Kapitel 4

Software Switch Testing Utility

The original plan for testing STPViz was to use cheap routers running *OpenWrt*[14] with STP enabled. This did not work for multiple reasons.

- *OpenWrt* routers handle the internal switch as one interface, which stops it from disabling single switch ports. This also makes their STP implementation useless, as broadcast storms will still occur when using multiple *OpenWrt* routers in a topology with loops.
- Through testing we found that the packets sent by the *OpenWrt* implementation are not recognized by our other devices. The other devices were a *TP-Link TL-SG2008 Gigabit Smart Switch* and an *ASUS RT-N56U Wireless Router*. These two devices would just forward the *OpenWrt* STP packets while still sending their own.
- We also found that *OpenWrt* just forwards received STP packets, without appending its own data. This alone would make it unusable for our purposes.

After making these discoveries, we tried *dd-wrt*[5] as an alternative. Unfortunately, it mostly shares the same behaviour with *OpenWrt* and we did not know whether or not we could make either of these operating systems work for this thesis. With two *dd-wrt* routers connected to each other (to create a loop) and STP enabled, broadcast storms did not occur. This was still true when connecting them on more than two ports. The other problems existed like in *OpenWrt*, making it unusable for this thesis as well.

In order to be able to stay within the time limit while still testing STPViz we decided to implement a software bridge capable of STP. It uses *pcap* to handle incoming packets and react to STP packets. Our *software-switch* can be run on any number of interfaces. However, it has not been tested on more than two. The source code for the *software-switch* can be found in its git[16]. It was written in C, in contrast to the *stp-tree-generator* which was written in C++.

4.1 Saving Switching Data

Listing 4.1 shows the variables and arrays we use for saving the switching data.

Listing 4.1: The Containers for Switching Data

```

1 pthread_mutex_t ifaceMutex;
2 int n;
3 char **names;
4 unsigned char **neighbours;
5 unsigned char **interfaces;
6 unsigned char *bridgeId;
7 unsigned char ***macTable;
8 int *macIndices;

```

The *names*, *neighbours* and *interfaces* arrays are arrays of `strings`. They store the local interface names (as returned by the `ifconfig` command), the connected MAC addresses and the local MAC addresses respectively. As basic C does not have strings, these are arrays of **char** arrays. In C, arrays are saved as pointers to the first element, with the other elements following sequentially after that. This makes *names*, *neighbours* and *interfaces* pointers to pointers. Because the *software-switch* can be run on multiple interfaces we need to keep track of the number of interfaces it is running on, which we do in *n*. To know which interface to forward a packet on, switches keep a so-called MAC table. We keep a list of MAC addresses for every interface, making *macTable* an array of arrays of **char** arrays. This is due to us saving MAC addresses as **char** arrays as well.

The *ifaceMutex* variable (mutex stands for mutual exclusion) is used to synchronize our different threads. We use two threads per interface. One to handle and forward incoming packets, as well as update our STP port states. The other one sends the STP packets every *helloTime* seconds. Sharing memory between *threads* (the C thread implementation we used) does not require any setup, as all heap memory is shared[12]. We need to synchronize these accesses to keep the different threads from corrupting data due to concurrent writes.

4.2 Handling STP Packets and Port States

The code that extracts data from the STP packets is identical to STPViz, due to the packets having the same format. Handling of this data is a lot different however. The code for this can be found in Listing 4.2. Parameters prefixed with *r* or *b* are gained from the incoming packet and stand for root and bridge (first hop) respectively. Values prefixed with *root* are for the root information currently saved in the bridge. *currentIndex* is the index of the interface the packet was received on.

Listing 4.2: Updating Port States

```

1 void updatePortStates(int currentIndex, unsigned char rPriority, unsigned
  char rExtension, unsigned char *rMac, unsigned int pathCost, unsigned
  short age, unsigned char bPriority, unsigned char bExtension){
2     pthread_mutex_lock(&ifaceMutex);
3
4     //check if the information on the ROOT port has changed
5     if(states[currentIndex] == ROOT){
6         if(compareBridges(rPriority, rExtension, rMac, rootPriority,
            rootExtension, root) != 0 || pathCost +portCost !=
            rootPathCost){

```

```

7         //force reset
8         for(int i=0; i<n; i++){
9             states[i] = DEDICATED;
10            messageAge = 0;
11        }
12    }
13 }
14
15 //check for a root change
16 if(compareBridges(rPriority, rExtension, rMac, rootPriority,
17    rootExtension, root) < 0 ||
18    (compareBridges(rPriority, rExtension, rMac, rootPriority,
19    rootExtension, root) == 0 && pathCost + portCost <
20    rootPathCost)){
21    memcpy(root, rMac, 6);
22    rootPriority = rPriority;
23    rootExtension = rExtension;
24    rootPathCost = pathCost + portCost;
25    messageAge = age+1;
26
27    for(int i=0; i<n; i++)
28        if(states[i] == ROOT)
29            states[i] = DEDICATED;
30
31    states[currentIndex] = ROOT;
32    sendTCN(currentIndex);
33 }
34
35 //check if we would be the correct root
36 if(compareBridges(priority, extension, bridgeId, rootPriority,
37    rootExtension, root) < 0){
38    memcpy(root, bridgeId, 6);
39    rootPriority = priority;
40    rootExtension = extension;
41    rootPathCost = 0;
42
43    for(int i=0; i<n; i++)
44        states[i] = DEDICATED;
45    messageAge = 0;
46 }
47
48 //if a port is in the BLOCKING state but shouldn't be, change it
49 if(states[currentIndex] == BLOCKING){
50     //if our root is the correct one, set the port to dedicated
51     if(compareBridges(rootPriority, rootExtension, root, rPriority,
52        rExtension, rMac) < 0)
53         states[currentIndex] = DEDICATED;

```



```

49
50     //if we have the same root, but have should be preferred, change
      to DEDICATED
51     if(compareBridges(rootPriority, rootExtension, root, rPriority,
      rExtension, rMac) == 0 &&
52         (rootPathCost < pathCost || (rootPathCost == pathCost &&
      compareBridges(priority, extension, bridgeId, bPriority,
      bExtension, neighbours[currentIndex]) < 0)))
53         states[currentIndex] = DEDICATED;
54 }
55
56 //if a port is DEDICATED but shouldn't be, change it
57 //only possibility should be same root different path cost
58 if(states[currentIndex] == DEDICATED){
59     //only change if the neighbour has the same root (smaller root is
      handled by root change, larger root is ignored -> stay
      DEDICATED)
60     //note that here we have a rootPathCost == pathCost (without
      adding the portcost). This is because here we are interested
      in whether or not they have the same distance to the root.
61     if(compareBridges(rPriority, rExtension, rMac, rootPriority,
      rootExtension, root) == 0)
62         //even then only change to BLOCKING if they have a shorter
      path or should be preferred
63         if(rootPathCost > pathCost || (rootPathCost == pathCost &&
      compareBridges(priority, extension, bridgeId, bPriority,
      bExtension, neighbours[currentIndex]) > 0))
64             states[currentIndex] = BLOCKING;
65 }
66
67 pthread_mutex_unlock(&ifaceMutex);
68 }

```

4.3 Handling Non-STP Packets

Non-STP packets are forwarded based on port states and knowledge of the target MAC address. We did not implement any special handling for broadcast messages. None of the packets that the bridge receives should ever come from a broadcast address, which means they should never be contained in the MAC table. As packets to unknown MAC addresses are broadcast by default, we do not need special behaviour.

Listing 4.3: Handling Non-STP Packets

```

1 //handle non-stp packets here
2 //skip packets originating from this interface
3 if(memcmp(ethh->h_source, interfaces[currentIndex], 6) == 0){

```

```

4     return;
5 }
6
7 //skip packets sent by another interface listener
8 for(int i=0; i<numPackages; i++)
9     if(sentPackages[i] != 0 && memcmp(sentPackages[i], bytes, header->len)
10        == 0)
11        return;
12
13 memcpy(sentPackages[packageIndex], bytes, header->len);
14 packageIndex = (packageIndex+1)%numPackages;
15
16 //add src mac to mac table
17 int found = 0;
18 for(int i=0; i<macTableSize; i++){
19     if(macTable[currentIndex][i] == NULL)
20         continue;
21     if(memcmp(macTable[currentIndex][i], ethh->h_source, 6) == 0){
22         found = 1;
23         break;
24     }
25 }
26 if(!found){
27     memcpy(macTable[currentIndex][macIndices[currentIndex]], ethh->
28         h_source, 6);
29     macIndices[currentIndex] = (macIndices[currentIndex] + 1)%macTableSize
30     ;
31 }
32
33 //check if the receiving interface is the intended target
34 if(memcmp(ethh->h_dest, interfaces[currentIndex], 6) == 0)
35     return;
36
37 //find right interface to send it on
38 int targetIndex = -1;
39 for(int i=0; i<n && targetIndex < 0; i++){
40     if(i==currentIndex)
41         continue;
42     for(int j=0; j<macTableSize; j++){
43         if(memcmp(macTable[i][j], ethh->h_dest, 6) == 0){
44             targetIndex = i;
45             break;
46         }
47     }
48 }
49
50 if(targetIndex < 0 || ethh->h_dest[0] & 1){

```

```

48     //if the packet was received on a blocking interface it is travelling
        sideways
49     //and needs to be stopped
50     if(states[currentIndex] == BLOCKING)
51         return;
52
53     //we don't know where to send it
54     //this sends the packet via all dedicated ports and the root if it is
        going up (coming from a dedicated port)
55     //and all dedicated ports if it is going down (coming from the root
        port)
56     for(int i=0; i<n; i++)
57         if(i!=currentIndex && states[i] != BLOCKING)
58             write(socks[i], bytes, header->len);
59 }else{
60     write(socks[targetIndex], bytes, header->len);
61 }

```

4.4 Installation and usage

This utility tool was written in C99[8]. While only one build target exists, we included a Makefile for convenience. The *pcap* development library is required to compile, and the regular *pcap* library is needed on launch. Following parameters can be set on launch:

- **-p** sets the STP priority (default: 0x80).
- **-e** sets the system id extension (default: 0).
- **-ms** sets the number of MAC addresses to store per interface (default: 30).

After these optional parameters the names of the interfaces to bridge are expected. The parameters used are printed to **stdout** for debugging purposes. Information on STP packets sent, including the respective port states is printed to **stdout** as well.

Kapitel 5

Testing

5.1 Setup

In order to ensure STPViz fulfills the requirements stated in Chapter 1 we performed several tests. To be able to check the results for correctness, we kept the testing network small and simple. The layout of the test setup can be seen in Figure 5.1. While a very basic network, it is sufficient for the tests described below, which will guarantee the required capabilities of STPViz. We changed the priority of bridge *C* to 61440 (the maximum) to ensure it would not be root. For bridges *A* and *B* we used the systemd id extension to ensure they would not be root when our *Root* bridge was connected. Nodes *A*, *B* and *C* were running *Wireshark* and STPViz. Additionally using *Wireshark* allowed us to check the actual packages involved in the testing process to monitor progress and note situations not handled correctly by STPViz. The server for STPViz was run on Node *C*. Test results were checked for their correctness by checking the report provided by STPViz against the layout the network actually has. This layout was deduced using the debug output from our running *software-switch* instances and the configuration of the hardware switches.

Actual Physical Setup

Due to our problems with *OpenWrt* and *dd-wrt* we had to use our *software-switch* tool for testing. We ran it on two laptops, which were bridges *A* and *B*. USB to Ethernet adapters were used to give the laptops an additional Ethernet port and make them usable. To reduce the amount of ports we needed, we ran STPViz instances on the laptops themselves instead of on a separate node. Unfortunately, the laptop replacing bridge *B* sometimes got confused which interface to use for the default route when all interfaces were connected. This stopped it from sending data to the server, which caused the server to remove all its old data. It is not an issue with anything we developed, but rather *NetworkManager* running on the laptop. The tests were not really affected, as whether bridge *A* was contained in the visualization was irrelevant in the cases where it happened. We decided not to set the timeout in the server to a large number to circumvent this, because this might have interfered with the actual test we wanted to perform.

In an actual scenario, where STPViz is running on a separate device, this is not an issue. Not even if some of the bridges in the network were running our *software-switch* utility, as

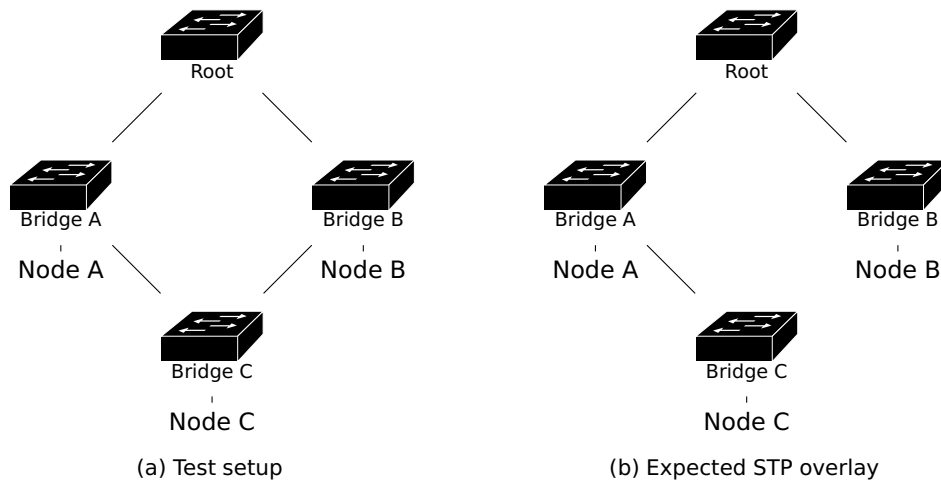


Abbildung 5.1: The expected physical and logical network topologies

basic switching was tested and is functional.

The devices we used have the following bridge IDs:

- **Root Bridge:** BC:AE:C5:EB:7D:B6
- **Bridge A:** 00:E0:7C:C8:57:E7
- **Bridge B:** 00:E1:00:00:0D:C4
- **Bridge C:** F4:F2:6D:7D:BF:BD

5.2 Tests

Plug and PlayTest

This test is meant to ensure the general bridge discovery ability of STPViz. It is designed to simulate the connection of nodes to an established network. Using a simple setup like shown in Figure 5.1a allows us to collect information about the whole network. If we were using less nodes than we had non-root bridges in the network, this would not be possible.

Performing the test

1. The layout shown in Figure 5.1a was established and all devices were started.
2. No STPViz instance was started yet, identification during the tree establishment is covered by Section 5.2: Tree Establishment Test.
3. We waited for the bridges to establish a stable spanning tree, checking the progress by observing the STP packets for the TC flag.

4. After the tree had stabilized we started the server on node *A*, as well as STPViz on all the nodes.
5. We waited for the STPViz instances to send their data to the server, which takes one *helloTime* plus the latency between the nodes and the server.
6. When all nodes had sent their data to the server we created a report to check it against the expected result.

Expected Result

With the limited size of the test network all bridges can and must be identified correctly for this test to be counted as successful. If the bridges in the network behave as they should, STPViz should not be able to gather information about connections between the bridges. Figure 5.2a shows the output we expect.

STPViz Output

All nodes are correctly identified. As the tree is already established when STPViz is started on all nodes, no information about connections between the bridges can be gathered. Figure 5.2 shows a comparison between expected output and STPViz output.

Tree Establishment Test

To test whether STPViz can handle bridges being added to the network during runtime, we start STPViz on all nodes before the establishment of the spanning tree, and check for correct identification afterwards.

Performing the test

1. The layout shown in Figure 5.1 was established and all devices were started.
2. Before enabling STP on all the bridges we started the server and STPViz instances.
3. After enabling STP on the bridges we waited for the tree to be established, again checking the progress via the TC flag.
4. Finally we checked the result for correctness.

Expected Result

Again all the bridges can and must be correctly identified. Additionally, this time the identification of bridge connections is possible, and the topology should therefore be fully identified.

STPViz Output

If bridge *D* is connected to bridges *B* and *C* before they are connected to bridge *A*, the tree can build up slowly and the clients will know about the connections. The resulting visualization can be seen in Figure 5.3

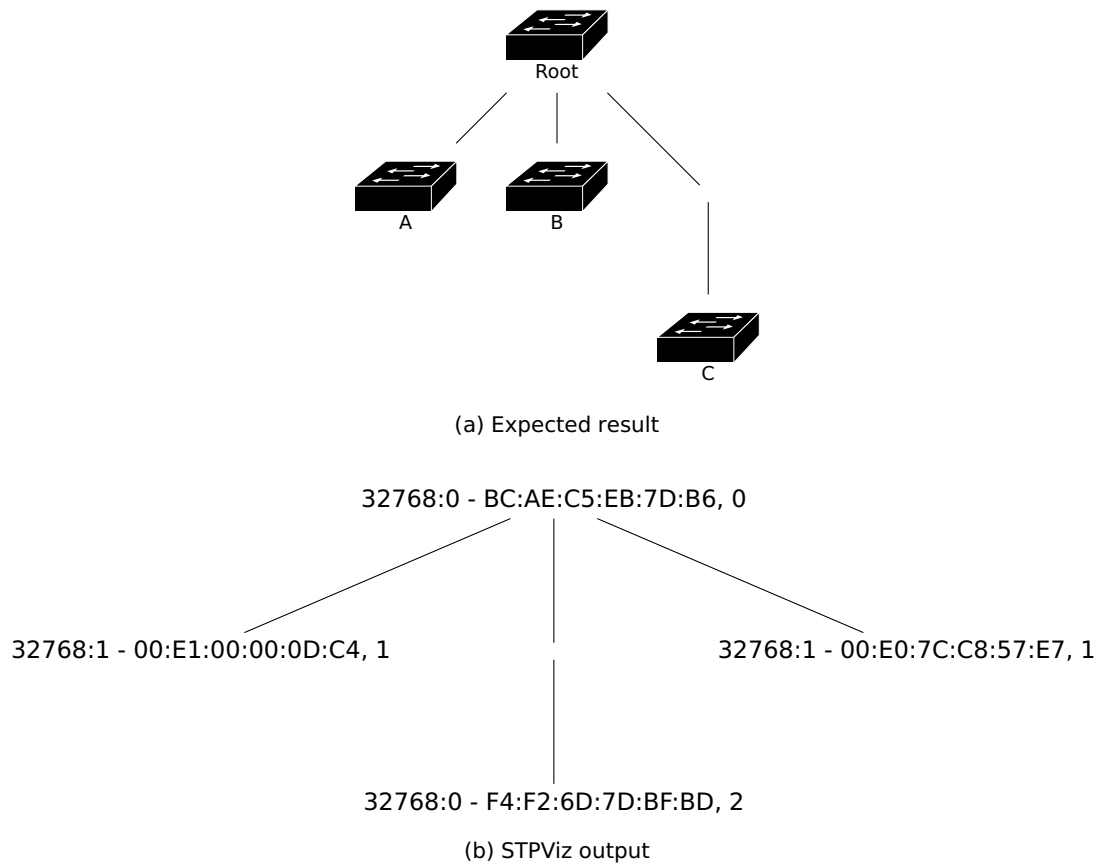


Abbildung 5.2: Expected output compared to STPViz output (Plug and PlayTest)

Bridge Removal Test

STPViz must also be able to react to bridges or clients dropping from a network. This test makes sure that capability is given.

Performing the test

1. We started all the nodes STPViz instances and waited for the spanning tree to be established and correctly identified (see the sections on the Usage Test 5.2 and the Tree Establishment Test 5.2).
2. After the tree was constructed we unplugged Node A and waited for the tree to stabilize.
3. Finally we checked the output of the report for correct identification of the smaller tree.

Expected Result

Bridge A must correctly be removed from the topology visualization. The expected topology is shown in Figure 5.4a.

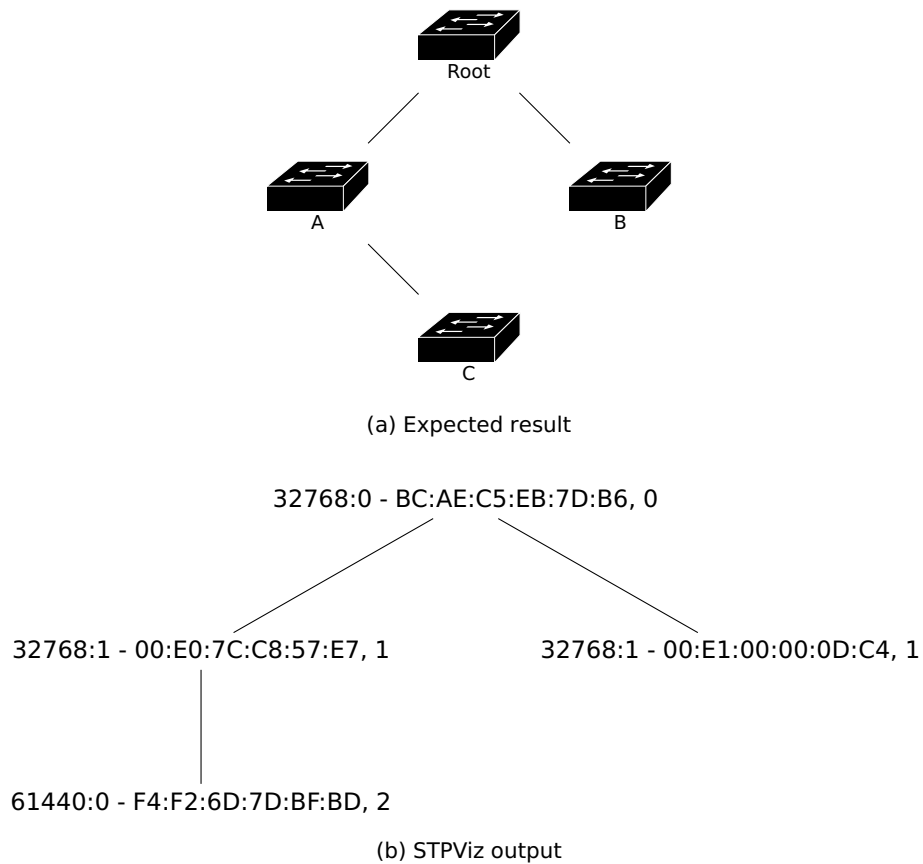


Abbildung 5.3: The result of the Tree Establishment Test

STPViz Output

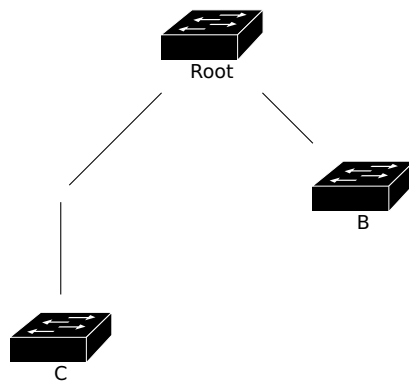
As visible in Figure 5.4, the bridge is correctly removed. This is one of the cases where the laptop running bridge *B* stopped sending data to the server. The removal of bridge *A* is still visible however, and so we deemed this test successful.

Slow Dynamic Change Test

Changes in the network topology must not be a problem for STPViz. By successfully testing for additions and removals from the network (Section 5.2: Plug and Playtest and Section 5.2: Removal Test) one could assume that changes can be handled as well, but caution demands that we test specifically for changes in the topology. Due to the rather complex nature of this test, Figure 5.5 shows the physical network topology after every step.

Performing the test

1. We set up all the nodes and waited for a stable tree to be generated by the bridges like before.



(a) The expected result

32768:0 - BC:AE:C5:EB:7D:B6, 0

61440:0 - F4:F2:6D:7D:BF:BD, 2

(b) The STPViz output

Abbildung 5.4: Expected output compared to STPViz output (Tree Removal Test)

2. We disconnected bridge A from the root, resulting in a topology like in Figure 5.5b.
3. The connection between bridges B and C was also severed, leaving bridges A and C in a disconnected subtree, as shown in Figure 5.5c.
4. We waited for the spanning tree to stabilize after the removal.
5. Bridges A and C were connected to bridge B yielding the physical topology shown in Figure 5.5d.
6. After the tree had stabilized we checked the report output for correctness.

Expected Result

Because we wait for the small subtree to stabilize, establishing bridge A as root, the subtree should be correctly identified, with the connection between the bridges. Bridge A should not be connected to this subtree in the output, as the Root bridge should be recognized immediately on reconnecting.

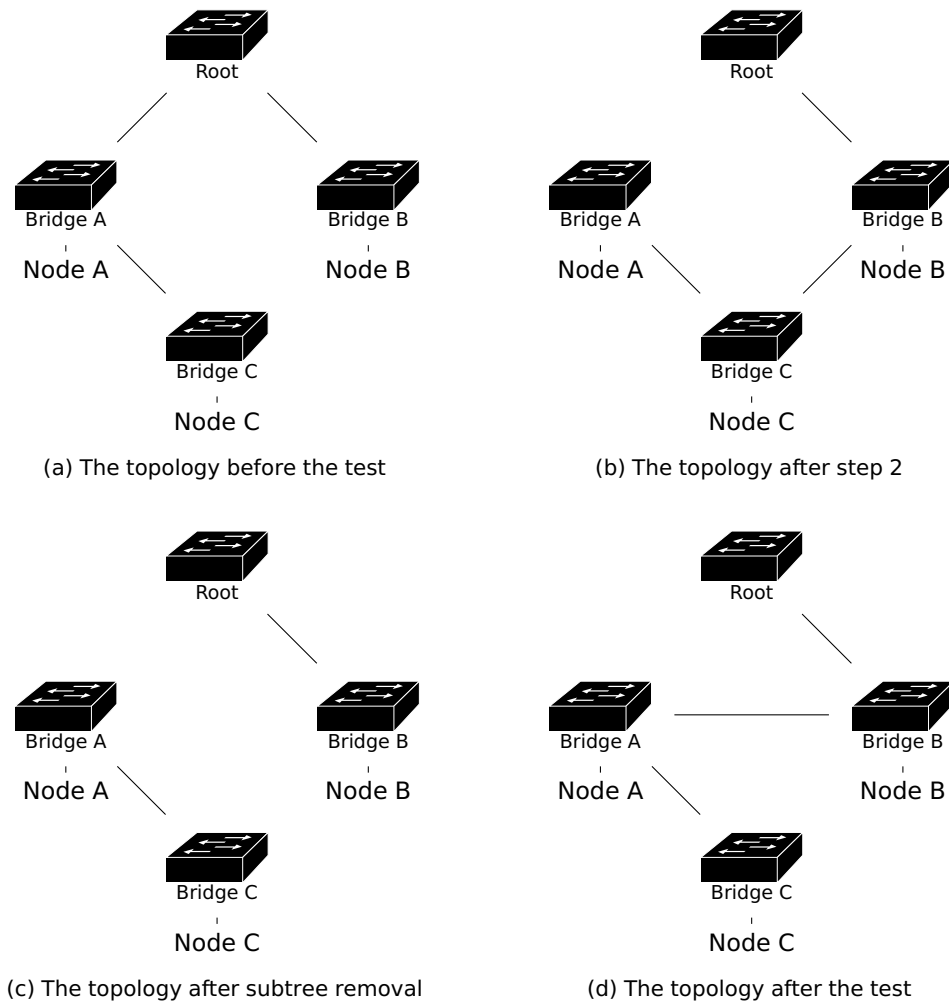


Abbildung 5.5: The physical topology during the Slow Dynamic Test

STPViz Output

During the test we checked the server output to see how disconnecting the two clients affected the output. We were surprised to see that it had not changed. After checking the *Wireshark* output, we realized that this was because no TC flag had been sent. Figure 5.6 shows the packet capture, including the changed packets. We marked the line after which the root identifier is updated. Where it should say *Conf. + TC* (meaning a configuration packet with set TC flag), a regular configuration packet was indicated.

As we realized during earlier testing, the TC flag is not always sent when a change occurs. This is why the client also looks for changes in message ages and the bridges in the STP packet. Due to the message age decreasing (one of the bridges is now root), the client will clear its data even without a TC flag. With the subtree stabilized, the client will prepend the *Root* bridge to A and C on reconnection.

66558	2511.5488967	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/0/bc:ae:c5:eb:7d:b6	Cost
66607	2513.5479415	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/0/bc:ae:c5:eb:7d:b6	Cost
66673	2515.5479364	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/0/bc:ae:c5:eb:7d:b6	Cost
66712	2517.6311279	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/0/bc:ae:c5:eb:7d:b6	Cost
66762	2519.6311078	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/0/bc:ae:c5:eb:7d:b6	Cost
66772	2521.2144855	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
66806	2521.7144133	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
66846	2523.7977492	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
66898	2525.8810913	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
66944	2527.9643999	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
66977	2530.0477598	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
67017	2532.1311868	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
67058	2534.2143872	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
67101	2536.2977871	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
67136	2538.3810508	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
67196	2540.4643712	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost
67233	2542.5477340	Tp-LinkT_7d:bf:bd	Spanning-tree-(for...	STP	64	Conf.	Root = 32768/1/00:e0:7c:c8:57:e7	Cost

Abbildung 5.6: The packet capture not showing a TC flag

The visualization acquired before executing the test is the same as in Figure 5.3, as we used that test as preparation. Unfortunately, bridge *B* stopped sending again. However, the test was a success, as the subtree was correctly identified and added to the large tree. The visualization we obtained from STPViz can be seen in Figure 5.7.

Fast Dynamic Change Test

Robustness to changes in the topology are a must for STPViz. It would however also be nice if STPViz were robust enough to handle topology changes without a spanning tree stabilization in between. The logical topology is the same as for the Slow Dynamic Change Test 5.2

Performing the test

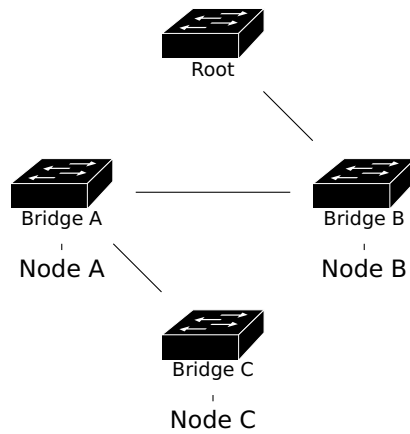
The steps required to perform this test are the same as for Section 5.2: Slow Dynamic Test, with one difference. This time we did not wait for the subtree to stabilize (after Figure 5.5c), but instead immediately plugged bridge *A* into bridge *B*.

Expected Result

We did not know what to expect. The nodes should be correctly identified, but we were unsure about the connections, as it was not clear how each bridge would update.

STPViz Output

For the first few seconds after reconnecting the subtree to bridge *B* (Figure 5.5d) nothing happened. After that, the bridges updated their information, which resulted in the same visualization as the Slow Dynamic Change Test. This means that our approach to keeping and saving bridge connection data is stable enough to withstand sudden topology changes. The output is identical to Figure 5.7.



(a) The expected result

32768:0 - BC:AE:C5:EB:7D:B6, 0

32768:1 - 00:E0:7C:C8:57:E7, 2

61440:0 - F4:F2:6D:7D:BF:BD, 3

(b) STPViz output

Abbildung 5.7: Expected output compared to STPViz output (Slow/Fast Dynamic Change Test)

Kapitel 6

Conclusion

The purpose of this thesis was to create a distributed, passive, low impact network visualization tool that relies solely on STP packets. With STPViz, we created a network visualization tool that does just that. It can extract information about bridges from STP packets, and under the right circumstances even make assumptions about bridge connections. When the established spanning tree is expanded, and a new root is prepended, STPViz expands its known topology with it. The inherent uncertainty of these assumptions is reduced by combining data obtained from distributed nodes. This means that the usefulness of STPViz increases greatly with the number of clients running on a network. We tested STPViz to the greatest extent possible with the limited hardware we had at our disposal. In order to perform these tests we had to develop *software-switch*, a utility that mimics STP capable bridges. It is a very basic and minimal tool, developed just far enough to make it usable for our tests. STPViz passed all the tests we designed (in static and dynamically changing networks). The tests showed that STPViz is fully capable of bridge discovery and handling a bridge's removal, as well as making assumptions about bridge connections on topology updates. In our tests, STPViz correctly filtered out incorrect assumptions in the server by comparing information from different clients. When confronted with sudden changes in the physical network topology, STPViz proved to be stable enough to handle them without discarding any correct assumptions made beforehand.

Unfortunately we did not find any other general method of checking assumed topologies. The lack of mechanics like a hop count on the Data Link layer keeps us from implementing something like a *traceroute*, which we could use to check our assumptions manually.

Future extensions of STPViz could include features to increase usability, like a GUI, as well as other output formats. STPViz could also be extended to include assumptions about actual cables by using the root path cost. Extending the functionality of STPViz to STP extension, as well as alternatives to STP (discussed in Section 2.2) would also be very helpful. An even better extension would be to detect and use whichever features the connected bridge is capable of dynamically at runtime. Additionally we should extend the parser to be capable of generating output from one or more *.pcapng* files without needing to launch a server and multiple clients beforehand. With more extensive testing and statistical analysis of the results, it is our opinion that the assumptions we make could be improved as well. As we realized during the Slow Dynamic Change Test (Section 5.2), TC flags are not always sent when they should be. Any future revision of this tool should consider finding another method of detecting network changes, to increase accuracy.

The *software-switch* utility could also be extended and tested more thoroughly to make it usable as a general purpose switching utility. For easier testing (e.g. with STPViz) it would be beneficial to find a way to create multiple virtual switches on a single device. This would greatly increase testing speed and reduce the amount of hardware required for testing large networking scenarios.

Other extensions of this thesis include work on the *OpenWrt* and *dd-wrt* projects in order to find a viable method to keep the STP implementation consistent with other hardware. Informing their respective communities of this issue and working together to find a solution would be a good start. Searching for ways to keep STP implementations consistent across devices and distributions would be a great benefit to the networking community.

Literaturverzeichnis

- [1] Richard A. Becker, Stephen G. Eick, and Allan R. Wilks. Visualizing Network Data. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):16–28, March 1995.
- [2] Ryan Blue, Cody Dunne, Adam Fuchs, Kyle King, and Aaron Schulman. *Visualizing Real-Time Network Resource Usage*, pages 119–135. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [3] brianwhelton. SolarWinds Thwack Forum. <https://thwack.solarwinds.com/ideas/3034>. Accessed: 2016-09-06.
- [4] Brocade. STP Parameters and Defaults. http://www.brocade.com/content/html/en/configuration-guide/FI_08030_L2/GUID-256DE2C6-C2DC-4DE0-A0C2-A96DDF0C3406.html. Accessed: 2016-09-23.
- [5] embeDD GmbH. dd-wrt. <https://dd-wrt.com>. Accessed: 2016-09-29.
- [6] IEEE Standards Associaton. IEEE 802.1aq Standard. <https://standards.ieee.org/findstds/standard/802.1aq-2012.html>. Accessed: 2016-09-28.
- [7] IEEE Standards Associaton. IEEE 802.1D Standard. <http://standards.ieee.org/getieee802/download/802.17a-2004.pdf>. Accessed: 2016-09-28.
- [8] International Organization for Standardization. ISO/IEC 9899:1999. http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=29237. Accessed: 2016-09-29.
- [9] L2Discover. L2Discover. <http://www.inf.u-szeged.hu/~bilickiv/l2discover/>. Accessed: 2016-09-06.
- [10] LiveAction. LiveNX Network Performance Management. <http://www.liveaction.com/>. Accessed: 2016-09-06.
- [11] LorientPro. LorientPro Network Management Software. <http://www.lorientpro.com/>. Accessed: 2016-09-06.
- [12] Kunle Olukotun. Shared Memory Parallel Programming, Pthread/OpenMP Examples, 2011. Stanford University, http://web.stanford.edu/~tayo/temp/aust_parallel/lect.02.ShMemPrg.pdf, Accessed: 2016-09-21.
- [13] open-source-parsers Github Community. A C++ Library for Interacting With JSON. <https://github.com/open-source-parsers/jsoncpp>. Accessed: 2016-09-06.

- [14] OpenWrt Community. OpenWrt. <https://openwrt.org/>. Accessed: 2016-09-29.
- [15] Radia Perlman. An Algorithm for Distributed Computation of a Spanningtree in an Extended LAN. *SIGCOMM Comput. Commun. Rev.*, 15(4):44–53, September 1985.
- [16] Alexander Schlögl. software-switch. <https://github.com/alxshine/software-switch>. Accessed: 2016-09-29.
- [17] Till Tantau. *The TikZ and PGF Packages*. Institut für Theoretische Informatik, Universität zu Lübeck, 1.18 edition, June 2007.
- [18] tcpdump & libpcap Community. tcpdump & libpcap. <http://www.tcpdump.org/>. Accessed: 2016-09-29.
- [19] The Linux Information Project. Broadcast Storm Definition. http://www.linfo.org/broadcast_storm.html. Accessed: 2016-09-09.
- [20] Krist Wongsuphasawat, Pornpat Artornsombudh, Bao Nguyen, and Justin McCann. Network Stack Diagnosis and Visualization Tool. In *Proceedings of the Symposium on Computer Human Interaction for the Management of Information Technology*, CHiMIT '09, pages 4:29–4:37, New York, NY, USA, 2009. ACM.