An algorithm for selection of the preferable root switch for the spanning-tree protocol in computer networks

P A Rahman

Department of Automated Technological and Informational Systems, Ufa State Petroleum Technological University, 2, October Avenue, Sterlitamak, 453118, Russia

E-mail: pavelar@yandex.ru

Abstract. This scientific paper deals with the computer networks based on the Ethernet technology and Spanning Tree Protocol (STP), as well as with the problem of selection of the root switch for the STP. The author offers an algorithm for selection of the preferable root switch for the STP based on the Floyd-Warshall algorithm for calculation of the shortest paths between all pairs of switches in a computer network. The algorithm for calculation of tree of the shortest paths for the given root switch is also overviewed. Finally, an example of the selection of the preferable root switch for the given network and calculation of tree of the shortest paths for the selected root switch is also presented.

1. Introduction

In present days the computer networks [1-2] have become an essential part of the modern enterprises. Computer networks are used both for data exchange between computers inside the enterprise and for communication with external networks of the Internet.

The computer networks of enterprises in most cases are based on technologies of the local area networks, such as Ethernet, and network switches, which support the Ethernet technology [3]. By default, the Ethernet technology can be used only in networks with the tree topology. The loops in the Ethernet networks are strictly prohibited because they cause rapid growth of amount of the Ethernet broadcast frames inside the network loops and this makes the network inoperable. On the other hand, the redundant physical links are often used for providing the fault-tolerance of local area networks, and they create loops in the network. Therefore, to resolve this problem, the specialized protocols are used, such as Spanning-Tree Protocol [4]. At the first step, the STP protocol selects one of the switches as the root switch. It then calculates a logical switching tree of the shortest paths between the root switch and other switches based on the physical topology of links between the switches. All physical links, which are not belonging to the calculated switching tree, are logically blocked. Next, in case of failure of switches or physical links, the STP calculates new logical tree, and this provides fault-tolerance of the local area network with redundant physical links.

By default, the STP automatically selects the root switch with the lowest ID (MAC-address of the switch), and this selection is most cases maybe not optimal from the viewpoint of disposition of the root switch in the network. On the other hand, in STP, it is possible to select the preferable root switch manually by using the particular STP parameter – switch priority. In this case, STP is a select switch with the lowest priority as the root. However, there is a scientific problem of how to calculate the preferable root switch in network with the given topology, and here we need a specialized algorithm.
Within the scope of scientific research in the field of computer networks [5, 6] the author developed a specialized algorithm for selection of the preferable root switch for the STP in computer networks based on algorithms of the graph theory [7-8] and discrete mathematics [9-10].

2. Automatic selection of the root switch in the Spanning-Tree Protocol

Let us discuss a simple fault-tolerant computer network with five switches and redundant topology of physical links. Figure 1 shows the physical topology of the computer network.

![Physical Topology of Computer Network](image1.png)

**Figure 1.** The physical topology of the computer network

In the Ethernet-based local area networks, all switches by default have the same STP priority, which is equal to 32768. Therefore, to select the root switch in a network, the STP only compares the MAC-addresses of switches and selects switch with the lowest MAC-address. Next, the STP calculates the tree of the shortest paths between the root switch and other switches using the cost information of physical links. By default, the cost of each physical link is obtained using the link bandwidth information and appropriate cost value defined for the Spanning-tree protocol (Table 1).

As a result, for the physical topology given above, the STP selects switch on the top of the network with the lowest MAC-address as the root switch and calculates logical switching tree of the shortest paths. Figure 2 shows the logical switching tree of the computer network.

All physical links, which are not belonging to the switching tree, are logically blocked.

![Logical Switching Tree of Computer Network](image2.png)

**Figure 2.** Logical switching tree of the computer network
To specify manually one of the switches as the preferable root switch we can set the high STP priority for this switch. The STP priority is an integer value in the range from 4096 to 32768 with the step of 4096. The value 4096 is considered as the highest priority and value 32768 is considered as the lowest priority. Moreover, we can specify one primary root and several backup root switches, using different STP priority values, such as 4096 for the primary root and 8192 for the backup root switches. Now let us discuss a specialized algorithm for selection of the preferable root switch.

3. A specialized algorithm for manual selection of the preferable root switch
Let the physical topology of computer network is given in the following form:

*n* is the number of switches in the physical topology of the given network.

\[ L = [[[i, j], w_{ij}]] \]

is the list of physical links between the switches and their costs. Each element of the list contains the indexes *i* and *j* of the linked switches and cost \( w_{ij} \) of the link.

The offered by the author algorithm for selection of the preferable root switch and calculation of the switching tree for the given physical topology of the network includes the following steps:

- Calculation of the adjacency matrix for the given physical topology of the network.
- Calculation of the matrix of lengths of the shortest paths and matrix of «next-switches» in the shortest paths between the switches by using the Floyd-Warshall algorithm.
- Selection of switch with the lowest length of the shortest path to the most remote switch.
- Calculation of the adjacency matrix for tree of the shortest paths for the selected root switch.

Let us discuss the steps of the specialized algorithm in detail.

At the first step, the adjacency matrix \( C \) for the given number of switches *n* and list \( L \) of physical links is formed using the following calculation scheme:

\[
C_{ij} = \begin{cases} 
  w_{ij}, & (i \neq j) \land ((([i, j], w_{ij}] \in L) \lor ([j, i], w_{ji}] \in L)); \\
  \infty, & (i \neq j) \land ((([i, j], w_{ij}] \not\in L) \land ([j, i], w_{ji}] \not\in L)).
\end{cases}
\] (1)

It should be noted, that for unlinked pairs of switches the cost is considered as infinite.

At the second step, the matrix \( T \) of lengths of the shortest paths and matrix \( H \) of «next-switches» in the shortest paths between the switches are calculated using the Floyd-Warshall algorithm and adjacency matrix \( C \).

At first, the matrices \( T \) and \( H \) are initialized using the following values:

\[
T_{ij}^{(0)} = C_{ij}; \quad H_{ij}^{(0)} = \begin{cases} 
  0, & C_{ij} = \infty; \\
  j, & C_{ij} \neq \infty.
\end{cases}
\] (2.1)

### Table 1. Default link costs for the Spanning-Tree protocol and Rapid STP

| Link bandwidth | Link cost for STP (IEEE 802.1d-1998) | Link cost for Rapid STP (IEEE 802.1w-2004) |
|----------------|--------------------------------------|------------------------------------------|
| 10 Mbit/s      | 100                                  | 200000                                   |
| 100 Mbit/s     | 19                                   | 20000                                    |
| 1 Gbit/s       | 4                                    | 2000                                     |
| 10 Gbit/s      | 2                                    | 2000                                     |
| 100 Gbit/s     | –                                    | 200                                      |
| 1 Tbit/s       | –                                    | 20                                       |
Next, the matrices $T$ and $H$ are recurrently calculated using the following scheme:

$$
T^{(k)}_{ij} = \begin{cases} 
T^{(k-1)}_{ij} + T^{(k-1)}_{ik}, & \{T^{(k-1)}_{ij} > T^{(k-1)}_{ik} + T^{(k-1)}_{kj}\} \land (i \neq k) \\
T^{(k-1)}_{ij}, & \{T^{(k-1)}_{ij} \neq \infty\} \land (j \neq k) \land (T^{(k-1)}_{ik} \neq \infty); \\
H^{(k)}_{ij} = \begin{cases} 
H^{(k-1)}_{ik}, & \{T^{(k-1)}_{ij} > T^{(k-1)}_{ik} + T^{(k-1)}_{kj}\} \land (i \neq k) \\
H^{(k-1)}_{ij}, & \{T^{(k-1)}_{ij} \neq \infty\} \land (j \neq k) \land (T^{(k-1)}_{ik} \neq \infty); \\
\text{else};
\end{cases}
\end{cases}
$$

$$
\Rightarrow \begin{cases} 
i, j = 1\ldots n; \\
t_j = T^{(n)}_{ij}; \\
H_{ij} = H^{(n)}_{ij}.
\end{cases}
\tag{2.2}
$$

Element $T_{ij}$ of the matrix $T$ shows the length of the shortest path between the switches $i$ and $j$.

Element $H_{ij}$ of the matrix $H$ shows the index of the next switch for the switch $i$ on its shortest path to the switch $j$.

At the third step, the preferable root switch $u$ is selected using the calculated matrix $T$ of lengths of the shortest paths between the switches and following selection scheme:

$$
\begin{cases} 
i = 1\ldots n; \\
S_i = \max_{j=1\ldots n} [T_{ij}];
\end{cases}
\qquad \text{At first, the maximums } S_i \text{ of rows in the matrix } T \text{ are calculated. Next, the row with the minimal value of } S_i \text{ is selected. The index } i^* \text{ of the selected row is considered as the preferable root switch } u.
\tag{3}
$$

It is essential to mention that in some cases there can be several rows with the same minimum value $S_i$. In such cases, there are several candidates for the role of the preferable root switch, and we can use any one of them as the «primary» root, and the rest of them use as the «backup» root switches for the Spanning-Tree Protocol.

At the fourth step, the adjacency matrix $D$ for tree of the shortest paths for the selected root switch $u$ is calculated using the calculated matrix $H$ of «next-switches» in the shortest paths between the switches and adjacency matrix $C$.

At first the matrix $D$ is initialized using the following values:

$$
\begin{cases} 
i = 1\ldots n; \\
j = 1\ldots n; \\
D_{ij} = \begin{cases} 
0, & i = j; \\
\infty, & i \neq j.
\end{cases}
\end{cases}
\tag{4.1}
$$

Next, all elements of the matrix $D$, which relate to the tree of the shortest paths for the selected root switch $u$, are iteratively calculated using the following scheme:

$$
\begin{cases} 
v = 1\ldots n; \\
q^{(0)} = u; \\
k = 1\ldots k^*; q^{(k-1)} \neq v; \\
q^{(k)} = H_{q^{(k-1)}q^*}; \\
D_{q^{(k-1)}q^*} = C_{q^{(k-1)}q^*}; \\
D_{q^{(k-1)}q^*} = C_{q^{(k-1-1)}q^*}.
\end{cases}
\tag{4.2}
$$

The matrix $D$ can be used for presenting graphically the resultant tree of the shortest paths for the selected root switch $u$. The tree of the shortest paths is also considered as the logical switching tree for the given network and selected root switch.
4. Example of the selection of the preferable root switch and calculation of the switching tree

A computer network, containing \( n = 9 \) switches, is given. The topology of the network is given the as a list of physical links between the switches and their costs:

\[
L = \{[[1, 2], 19], [[1, 5], 19], [[1, 6], 4], [[2, 6], 4], [[2, 7], 4], [[3, 7], 4], [[3, 8], 4], [[4, 8], 4], [[4, 9], 4], [[5, 6], 4], [[5, 9], 4], [[6, 7], 2], [[6, 9], 2], [[7, 8], 2], [[7, 9], 2], [[8, 9], 2]\}.

Figure 3 shows the physical topology of the given network.

![Figure 3](image)

**Figure 3.** The physical topology of the given network

At the first step, we calculate the adjacency matrix \( C \) for the given number of switches \( n \) and list \( L \) of physical links between the switches and their costs:

\[
C_{ij} =
\begin{array}{cccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
1 & 0 & 19 & \infty & \infty & 19 & 4 & \infty & \infty & \infty \\
2 & 19 & 0 & \infty & \infty & \infty & 4 & \infty & \infty & \infty \\
3 & \infty & \infty & 0 & \infty & \infty & \infty & 4 & 4 & \infty \\
4 & \infty & \infty & 0 & \infty & \infty & \infty & 4 & 4 & \infty \\
5 & \infty & \infty & \infty & 0 & \infty & \infty & \infty & \infty & \infty \\
6 & 4 & 4 & \infty & \infty & 4 & 0 & 2 & \infty & \infty \\
7 & \infty & 4 & 4 & \infty & 2 & 0 & 2 & 2 & \infty \\
8 & \infty & \infty & 4 & 4 & \infty & 2 & 0 & 2 & \infty \\
9 & \infty & \infty & \infty & 4 & 4 & 2 & 2 & 2 & \infty \\
\end{array}
\]

At the second step, we calculate the matrix \( T \) of lengths of the shortest paths and matrix \( H \) of «next-switches» in the shortest paths between the switches:
At the third step, we calculate the maximums of rows in the matrix $T$ and then select the rows with minimal values of the calculated maximums:

$$
T_{ij} = \begin{bmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
1 & 0 & 8 & 10 & 10 & 8 & 4 & 6 & 8 & 6 \\
2 & 8 & 0 & 8 & 10 & 8 & 4 & 4 & 6 & 6 \\
3 & 10 & 8 & 0 & 8 & 10 & 6 & 4 & 4 & 6 \\
4 & 10 & 10 & 8 & 0 & 8 & 6 & 6 & 4 & 4 \\
5 & 8 & 8 & 10 & 8 & 0 & 4 & 6 & 6 & 4 \\
6 & 4 & 4 & 6 & 6 & 4 & 0 & 2 & 4 & 2 \\
7 & 6 & 4 & 4 & 6 & 6 & 2 & 0 & 2 & 2 \\
8 & 8 & 6 & 4 & 4 & 6 & 4 & 2 & 0 & 2 \\
9 & 6 & 6 & 6 & 4 & 4 & 2 & 2 & 2 & 0 \\
\end{bmatrix}
$$

$$
H_{ij} = \begin{bmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
1 & 6 & 6 & 6 & 6 & 6 & 6 & 6 & 6 \\
2 & 6 & 2 & 7 & 7 & 6 & 6 & 7 & 7 \\
3 & 7 & 7 & 3 & 8 & 7 & 7 & 7 & 8 & 7 \\
4 & 9 & 8 & 8 & 4 & 9 & 9 & 8 & 8 & 9 \\
5 & 6 & 6 & 9 & 5 & 6 & 6 & 9 & 9 \\
6 & 1 & 2 & 7 & 9 & 5 & 6 & 7 & 7 & 9 \\
7 & 6 & 2 & 3 & 8 & 6 & 6 & 7 & 8 & 9 \\
8 & 7 & 7 & 3 & 4 & 9 & 7 & 7 & 8 & 9 \\
9 & 6 & 6 & 7 & 4 & 5 & 6 & 7 & 8 & 9 \\
\end{bmatrix}
$$

As we can see, there are three rows 6, 7 and 9 with the minimal value of the row maximums. So, let us select the switch 6 as the preferable primary root switch, and the switches 7 and 9 consider as the backup root switches.

Finally, at the fourth step, we calculate the adjacency matrix $D$ for the tree of the shortest paths for the selected primary root switch $u = 6$:

$$
D_{ij} = \begin{bmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
1 & 0 & \infty & \infty & \infty & \infty & 4 & \infty & \infty & \infty \\
2 & \infty & 0 & \infty & \infty & \infty & 4 & \infty & \infty & \infty \\
3 & \infty & \infty & 0 & \infty & \infty & \infty & 4 & \infty & \infty \\
4 & \infty & \infty & \infty & 0 & \infty & \infty & \infty & 4 & \infty \\
5 & \infty & \infty & \infty & \infty & 0 & \infty & \infty & \infty & \infty \\
6 & \infty & 4 & \infty & \infty & 4 & 0 & 2 & \infty & 2 \\
7 & \infty & \infty & 4 & \infty & \infty & 2 & 0 & 2 & \infty \\
8 & \infty & \infty & \infty & \infty & \infty & \infty & 2 & 0 & \infty \\
9 & \infty & \infty & \infty & 4 & \infty & 2 & \infty & \infty & 0 \\
\end{bmatrix}
$$
The tree of the shortest paths is considered as the logical switching tree for the given network. Figure 4 graphically shows the logical switching tree for the selected primary root switch $u = 6$.

**Figure 4.** Logical switching tree for the selected primary root switch 6.

5. Conclusion

Thus, within the scope of this scientific paper, the author offered a specialized algorithm for selection of the preferable root switch for the Spanning-Tree Protocol based on the Floyd-Warshall algorithm for calculation of the shortest paths between all pairs of switches in a computer network. An example of the selection of the preferable root switch for the given network and calculation of tree of the shortest paths for the selected root switch is also presented.

The modern computer networks often use a set of virtual local area networks (VLANs) over the physical networks, and there are several modifications of the STP, such as Per-VLAN Spanning Tree Protocol (PVSTP) and Multiple Spanning-Tree (MST). These protocols allow specifying the preferable root switches separately for the different VLANs or groups of VLANs to provide high performance and reliability of the computer network. Therefore, in the case when we have several preferable root switches, we can use each of the preferable root switches as the primary root for one group of VLANs and as the backup root for other groups of VLANs.

Acknowledgements

The author is grateful to chief network administrator S.N. Khorkov from Moscow Power Engineering Institute for valuable pieces of advice in the field of computer network technologies.

References

[1] Peterson L L and Davie B S 2011 Computer Networks: A system approach (Morgan Kaufmann Publishers Inc.)
[2] Rybaczyk P 2004 Cisco Network Design Solutions for Small-Medium Businesses (Cisco Press)
[3] Spurgeon C E and Zimmerman J 2013 Ethernet Switches (O’Reilly)
[4] Seifert R and Edwards J 2008 The all-new Switch Book: The Complete Guide to LAN Switching Technology (Wiley Computer Publishing)
[5] Rahman P A and Bobkova E Yu 2017 J. Phys.: Conf. Ser. 803 012124
[6] Rahman P A 2018 J. Phys.: Conf. Ser. 1015 022016
[7] Zykov A A 1987 Basis of the graph theory (Moscow: Nauka)
[8] Kasyanov V N and Evstigneev V A 2003 Graphs in programming: processing, visualization and application (Saint-Petersburg: BHV-Petersburg)
[9] Gallier J 2011 Discrete Mathematics (New York: Springer-Verlag)
[10] Novikov F A 2009 Discrete mathematics for programmers (Saint-Petersburg: Piter)