Optimized Time Convergence in Switching And Routing.

B.G.Prasanthi†, Dr. T.Bhaskar reddy††
† S.R.N. Adarsh college / Asst.Professor & Head of the Department of Computer Science, Bangalore, India

†† S.K.University / Associate Professor , S.K.University,Anatapur,India.

Abstract
The wide deployment of Layer 2/Layer 3 switches in the wiring closet and distribution layers of the campus and enterprise networks has cause lot of dependency on the Spanning Tree Protocol to ensure loop free networks. To define, spanning tree is a protocol designed to eliminate loops in a topology having redundant links.
This paper tries to explore the trivial and widely used STP to introduce three enhancements, which can help in improving performance of the network in terms of convergence time ,route ability , better link utilization.
In the growing LAN network to control the redundant link failures in layer2 and to over come the issue faced by administrators in configuring different types of protocols.
Keywords:
STP, chassis, channeling, Zones, VLAB ,Switch, convergence .

1. Introduction and Spanning Tree Operation

The process of forwarding a frame from a source to the direction of its destination is a two-step process. First step is learning, and the second is forwarding of the frame. If the destination address is not known, the frame will be forwarded to all ports (except the port from where it arrived).
Spanning tree is needed in networks with redundancy, where undesirable loops can form.
To provide path redundancy, STP defines a tree that spans all switches in an extended network. STP forces certain redundant data paths into a standby (blocked) state. If one network segment in the Spanning Tree Protocol becomes unreachable, or if Spanning Tree Protocol costs change, the spanning tree algorithm reconfigures the spanning-tree topology and reestablishes the link by activating the standby path.
Many routing approaches have been proposed over the past several years. There are intricacies involved in the election of the root.[2]change of topology is also a pathetic situation which has to be taken in to consideration.[3]composed each net in to a set of two terminal nets and related them onto minimum distance paths.[4] sequential routing is allowed ,which purely takes place in flooding which increases the redundancy.[6] formulates the routing problem as an interval packing problem.[7][8] modeled for the construction of routing trees for the nets.[10] used to find the approximations and the complexity of representing a graph.[11] TRACER values has been considered ,since it is one of the efficient algorithm in net and path delays till now.
In this paper , we present a new novel routing algorithm .The key features of our approach are to avoid redundancy and to reduce net ,path, routability delays.
The results show that our routing solutions are very effective in reducing the net/path delays when compared to the above considered

2. Problem definition and conceptual solution

1) The Spanning tree protocol ensures a loop free network, but at a cost. The cost is that the redundant link will be blocked, and no traffic can flow through it. When network users buy a chassis being charged on a per port basis, this becomes considerable cost. It could be more serious for core or central layer 2 switches as it is important to have redundant links between such chassis. Consider fig 1, where there are two paths existing for traffic to flow from B to C. With STP active, only path via root will be open.

Although loop is avoided, this clearly has two disadvantages.

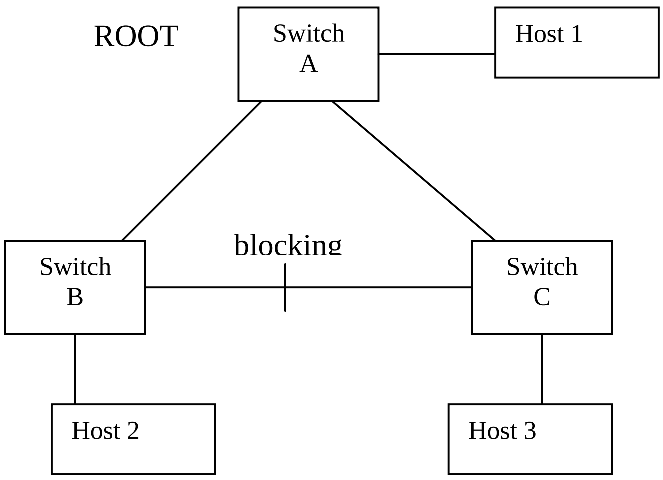


Figure.1

Manuscript received December 5, 2009
Manuscript revised December 20, 2009

1) Assume 1 Gb links connecting each of the switches, from B to C, effectively there is 2 Gbps bandwidth. But only 1Gb is practically available.
2) To reach B from C, the shorter path is directly from B to C, but being blocked, the path B-A-C has to be taken.

Solution

1) As solution, we introduce the concept of channeling among such redundant links. This is to be configured by the administrator, preferably over the high traffic zones, backbone switches. The channel will be formed typically between switches B and C non root devices here and will tell the software that there exists a redundant path, one directly and another via the root, A.

The administrator can either configure a traffic threshold, beyond which he might want the switch to use the redundant path (B to C).

The channel will consider the link B to C and B-A-C as one bundle, sharing common properties.

Alternative solution used in field now is to make the switches A, B and C as roots for different VLANs, so that different sets of links become as blocking and hence load balancing to an extent is achieved. But then if one VLAN is known to send more traffic than other or if we have just 1 or 2 VLANs, the problem is not effectively solved.
2) For networks having huge number or layer 2 devices, the convergence time is often not acceptable, just for any minor topology change. For instance, consider a network as in Figure 2, for a change/toggle in link, say X (marked in fig) the TCN exchange and then stabilizing the netwrk to be safe fortrafficforwarding would take atleast 30-60 secs.

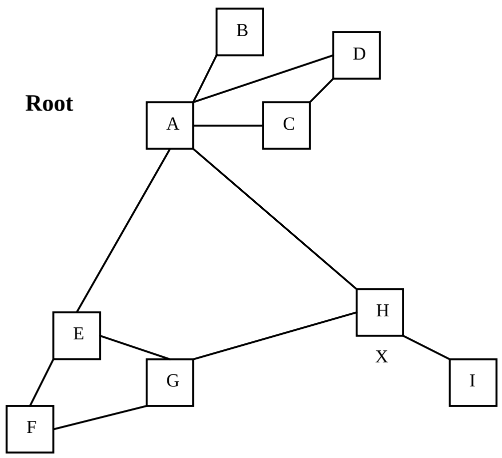


Figure 2

solution

To solve such issues in larger networks, we introduce the concept of division into zones or areas. This is similar to what routing protocols like OSPF use. The administrator has the task of dividing the network to zones. Each zone will have a primary root based on selection criteria similar to selecting root in a combined network.

Consider a zone to be comprising of switches E, F and G. Assume E is the root here. E now has the responsibility of talking (directly or indirectly) to the outside world. What we achieve out of this is that topology changes or fluctuations within a zone, like say switches A,B,C and D will be contained and resolved without putting the whole stream of switches downstream on discarding or blocking state.

3. Problem definition and conceptual solution

In the growing LAN network to control the redundant link failures in layer2 and to over come the issue faced by administrators in configuring different types of protocols.

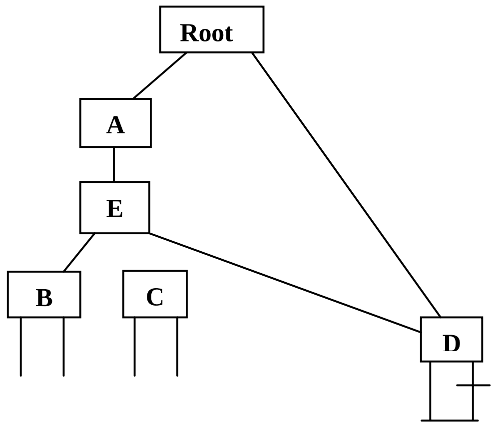


Figure 3

As there are four stages a node has to pass on from source to the destination in sending the packets. They are blocked state, listening state ,learning and forwarding states. So instead of all the nodes passing through four stages, it is time consuming so in order to fasten up the data sending ,the optimized solution is provided for the above fixed network scenario.

Switch B and switch C can be configured which allows the switch port to begin forwarding as soon as the end system is connected bypassing the listening and learning states and eliminating up to 30 seconds of delay before the end system can begin sending and receiving traffic.

Switch A and Switch D can be configured on access switch which is directly connected to root bridge and it is effective only when the configured switch has some non self looped blocked pairs.

Switch E can be configured with a switch which is indirectly connected to root port to know if nay indirect link failure happens for a root switch.

Merging step: This step tries to merge the connections in different connection/switch blocks to maximize resource sharing .The sharing of tracks will lead to a reduction in the number of switches used which would then directly contribute to the reduction in the roiuting delay of the net. Here is the algorithm as follows.

An algorithm to avoid path redundancy :

Set the nodes according to the path criticality estimated using the measure

For each set n_i in the list

- Find the edges in the net according to edge. path pi.

If(edge. Path pi is not already existing)

- Perform shortest path from signal source.
- Select the k-1 shortest two terminal nets whose route match with the signal flow of ni.

Else

- Assign edge-path pi to particular pair of terminals
- Construct a minimum spanning tree
- Generate two terminal net for each edge of the spanning tree.

End if;

Endfor.

An algorithm for root selection

- Construct a resource Graph G for the routes of subnets obtained from the above redundancy step.
- Determine the enumeration of each node by implicit technique.
- Find the root of the net by reporting the chromatic number.

An algorithm for merging

Set the nets according to the routing criticality;

Repeat

For each net n_i in the list

- Construct a resource graph G for n_i

Repeat

- From the source terminal node toward the sink.
- Find the pair of nodes that can share a track

If(there is a pair)

- Remove the flow between the nodes.
- Update the table.

End if

Until

End for;

Until(no reduction)

Detailed solution

Spanning tree channel:

Given a network as shown in fig 1, or say more complex networks as well like fig.3, we visualize the network as a channel, with one direct link and another virtual link. See fig 4.

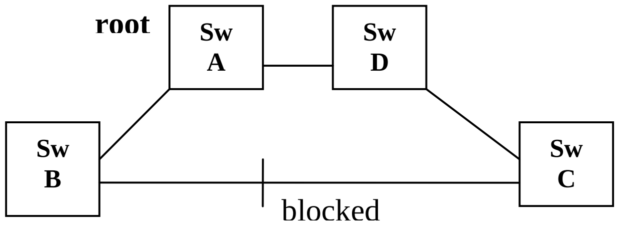


Figure 4

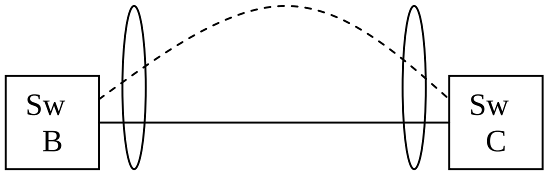


Figure 5

Lets consider a three DUT setup to see the detailed working, like fig 1.

Here, the administrator configures ports between non-root switches as parts of channel.

Assume there was no spanning tree. Host 1 sends traffic. Switch A will flood the packets to B and C. If the destination is known to B or C, it will be accordingly forwarded instead of flooding. But in the process, the address MAC 1 will be learnt at the port through which it arrived at B and C. If the destination is not reached, the frames will further be flooded, and B will receive it from C as well (and vice versa). So B, for example, now has 2 entries for host 1, reachable via Switch A and Switch C, and packets will this way get flooded again.

To prevent this, spanning tree blocks the link between B and C. Now, let us get channeling solution into the picture.

With channeling, the switch software knows that it can expect two ways to reach a particular destination via spanning tree channel ports. This, the switch has to effectively use by load balancing. The blocking port here maintains a new state called 'Blocked Channel'. The main concept is, after the learning phase, the traffic is not forwarded immediately by these ports. The MAC is learnt and stored.

As next step, all traffic received from other ports of B will be forwarded to both, C as well as towards root. C's responsibility is that, it will not forward any packets received at the 'blocked channel' port to other port of the channel. Any traffic addressed to destination MACs learnt in non-channel port will be forwarded by C. These MACs will be advertised by C in a separate format to B, mentioning that these are the MACs which can be reached directly via C. Directly here means, not via other channel port. In fig 5, B3 and B4 are directly reachable by B, as seen by C.

This way, we achieve the goal of load balancing and reaching via shortest path when possible.

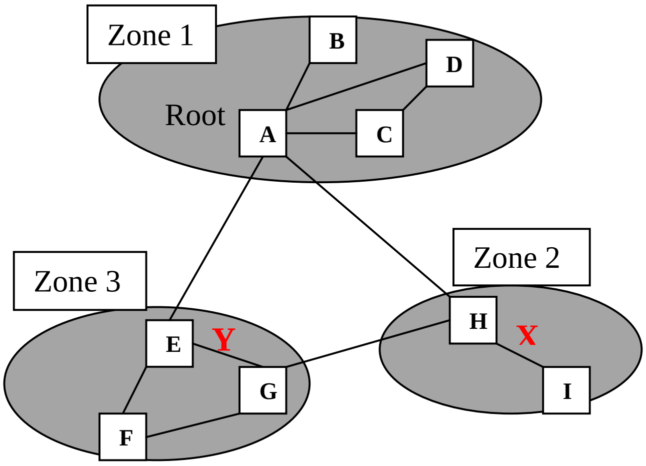


Figure 6

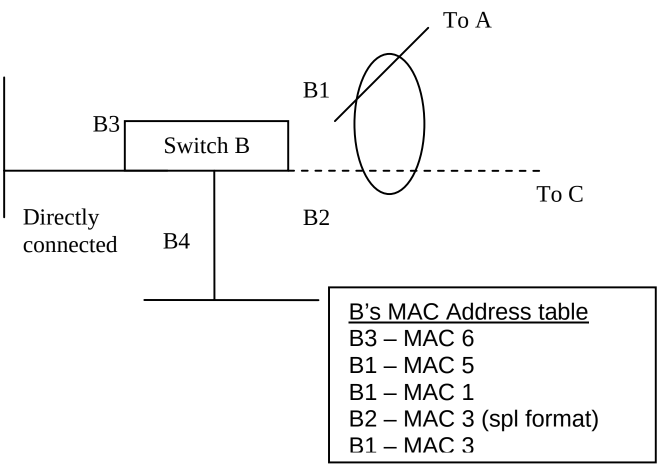


Figure 7

Taking examples, assume that Sw A floods traffic from say, MAC 5 towards B and C. Both B and C will check there CAM entries (both, blocked channel reachable (special format) MACs as well as its connected ones). See fig 5

If C does not advertise this as directly reachable while learning, it will not forward the traffic to C.

Now, assume that MAC 3 is directly connected to C. This is learnt by B using special format frames. Packets from ports B3 or B4 to MAC 3 will be load balanced by B and forwarded via B1 as well as B2. C will now entertain both, i.e. frames received from root as well as from blocked channel port.

Limitation:
This concept however has the limitation that the administrator has to have the additional configuration over the STP blocked ports, and the switch has the overhead of having two sets of MAC addresses.

4. Spanning tree zones

For this second level of optimization, consider the fig 2 as below, divided into zones.

The administrator configures the topology into zones. Spanning tree now elects roots within the zones.

Assume that in zone 1, A becomes the root and E in zone 3. Now, A only has to talk to roots of other zones. Among the roots, again one is elected as primary roots, which as per normal spanning tree BPDU exchange will be the lowest MAC. This way the additional overhead is less. Now, A sees zone 2 and zone 3 as just one device. Spanning tree convergence will result in the port between zone 3 and zone 2 to be blocked, ie E to H, via G. Switches like G here should act as transparent to BPDU exchange between zone roots.

This done, now to reap the benefits! Consider a link breakage as in Y (fig 6) which is in forwarding mode among E, F and G. Intra zone spanning tree will immediately activate or move the forwarding the link via F. All this will be transparent to zones 1 and 2, and with less number of switches per zone, convergence is a matter of milliseconds. Similar for a case of topology change.

Experimental Results:

We have implemented the proposed routing algorithm in C-Programming. We have tested our algorithm on a set of 50 nodes with 14 routers. We explore the tradeoff between the performance of routing.

Table 1:Net Delay

Circuits	SEGA	TRACER	ours
2 bit alu	110	354	78
4 bit alu	163	458	126
apex 7	78	147	52
term 1	70	113	39

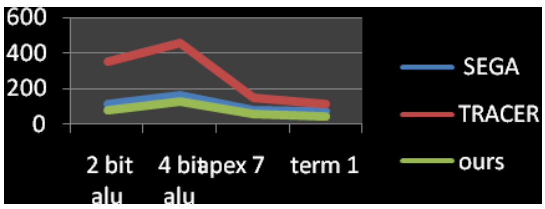


Table 2:Path delay

Circuits	SEGA	TRACER	ours
2 bit alu	947	714	618
4 bit alu	1392	1037	1444
apex 7	339	299	320
term 1	129	136	131

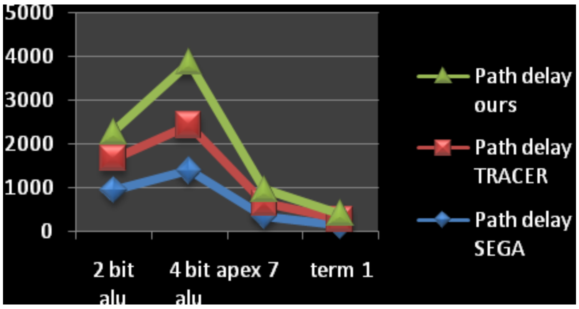
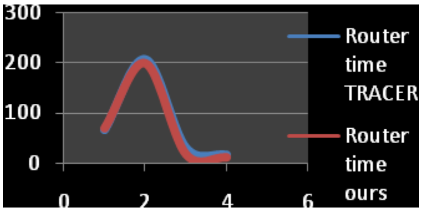


Table3 : Router time

Circuits	TRACER	ours
2 bit alu	66	70
4 bit alu	207	200
apex 7	30	18
term 1	16	12



Conclusion:

With the introduction of the concept of Spanning Tree Channels, the backbone switches can be configured to make use of the links which are added for redundancy, but are not used when primary link is up. This increases load balancing and overall network speeds.

The Spanning Tree Zone concept is mainly for huge clusters of switches which have spanning tree configured on them. These take high convergence times, even to address a link failure or topology change at one end. By dividing into zones & with the advanced features for the switches , the convergence time and recalculation overhead is considerably reduced and is not felt across the network . Extensive experimental data showed that the proposed routing algorithm is very effective in improving the overall performance of the design as well as routability.

References:

[1] King shaum lui, Whay Choui Lee , “A transparent spanning tree bridge protocol with alternate routing” - SIGCOMM Computer Communication,Vol 32 , p-33-46.
[2] L. Bosack and C. Hedrick. “ Problems in Large LANs”. IEEE Network Magazine, 2(1), p-22-28 ,Jan.
[3] Duato , José Serrano, “A New Transparent Bridge Protocol for LAN Internetworking using Topologies with Active Loops”, Proceedings of the 1998 International Conference on Parallel Processing, p.295-303, August 10-14, 1998.
[4] King Lui , Whay Chiou Lee, “ Spanning Tree Alternate Routing Bridge Protocol”, University of Illinois at Urbana-Champaign, Champaign, IL.
[5] Roy Friedman, “Efficient route discovery in hybrid networks” –Technion Institute of technology,p-69-75,2008.
[6] Michael Elkin ,”A faster distributed protocol for constructing a minimum spanning tree” , Journal of Computer and System Sciences, Vol 72 , P 1282-1308,2006.
[7] Hussein Hojjat, Hootan Nakhost, Marjan Sirjani “ Formal verification of Spanning tree protocol “ - Electronic Notes in Theoretical Computer Science, Vol159,p-139-154,2006.
[8] N. Linge, E. Ball, R. Tasker ,” A Bridge protocol for creating a spanning tree topology with in an IEEE 802 extended LAN environment” -Computer Networks and ISDN Systems , Vol 13,p-323-332..
[9] Y.-S. Lee and A. C.-H. Wu, "A performance and routability-driven router for FPGA's considering path delays," IEEE TCAD, Feb. 1997.
[10] A router for symmetrical FPGA based on excat routing density evaluation,”IEEE CSI vol53, no7,july 2004.
[11] Understanding and Designing Networks using Spanning Tree and UplinkFast Groups
[12] Ether channel concepts

***B.G.Prasanthi ,M.Sc.M.tech.,M.phil (CS)* from Sri Rama krishna PG college** ,Nandyal and M.phil from algappa university and doing Ph.D in computer science at S.K.University Anantaput. she has 8 years of teaching experience and 2 years of industry. She has published many national ,international papers in wireless networks, routing techniques She is currently working as Asst.Professor & head of the department, Computer Science in S.R.N.Adarsh College,Bangalore.

Dr.T.Bhaskar Reddy,M.Sc. (CS) Ph.D. Associate .Professor, Dept. of CST, Sri Krishnadevaraya University, Anantapur.

He received his M.Sc. and Ph.D. from Sri Krishnadevaraya University from 1997 to 2006 respectively. He has joined S.K University in 1989 and Presently working as Associate Professor in Department of Computer Science & Technology. He has published research papers in International, national journals and presented papers in international conferences also. His research areas of interest include Image Processing, Computer Graphics, , Network Security and Data Mining.