Abstract—The Ethernet frame goes round and round when an Ethernet network includes loop links. To solve this problem, the spanning tree protocol (STP) [1] has been proposed for constructing a virtual tree network on an Ethernet network when the Ethernet network includes loop links. The STP is implemented on various Ethernet switches. The STP suspends some physical links to construct the tree network. These suspended physical links stand by, and they are used for backup when there is trouble on a tree network link. Therefore, all traffic is concentrated on the virtual tree network. Even if traffic is jammed on the network, the suspended physical links are usually still not used.

Therefore, in this work, we describe a method that uses Ethernet Bypass Nodes (EBN) so that only communications between specific nodes can pass through the physical links suspended by STP as a bypass to improve communication quality.

Index Terms—Ethernet, Load Balancing, STP, NAT, Routing

I. INTRODUCTION

On Ethernet networks, the spanning tree protocol (STP) [1] is widely used for constructing virtual tree networks that aim to prevent Ethernet frame loops. The STP constructs the tree network topology on the basis of link bandwidth, Ethernet switch ID (also known as bridge ID), and each Ethernet switch’s port ID. Network operators can also set an optional Ethernet switch ID and then, construct a virtual tree topology that is the same as those designed on Ethernet networks.

However, the STP constructs only one tree network at a time on an Ethernet network. The STP suspends some physical links to construct a virtual tree network. The suspended links are used for reconstructing a new virtual tree network when any link is down on a current virtual tree network. However, the suspended links are not used for avoiding congestion even though communication between several specific nodes could avoid congestion by using them. As shown in above example, even though there are physical links in addition to a virtual tree network, all communications traffic is concentrated on the virtual tree, because the STP never constructs plural tree networks at the same time on an Ethernet network.

Therefore, in this research, we describe an Ethernet bypass method that aims to improve communication quality by traffic dispersion. Using this Ethernet bypass, only selected communications between specific nodes can pass through the STP’s suspended physical links that are not used for constructing a virtual tree network.

II. BACKGROUND

Early Ethernet technology had bus and tree topology, not the loop network topology, because the Ethernet technology was for bus type networks. Therefore, when there is loop network topology on an Ethernet network, a broadcast Ethernet frame rounds the Ethernet network without expiring. The Ethernet frame round not only wastes bandwidth but also causes incorrect learning of the MAC address table on the Ethernet switch. Moreover, Ethernet broadcast communication is required on an Ethernet network, because the address resolution protocol (ARP) requires Ethernet broadcasting for resolving from the IP address to MAC address.

Therefore, the spanning tree protocol (STP) [1] was proposed to prevent the Ethernet frame round. The STP constructs a virtual tree network as a loop-free network by turning some of the links that cause loop topology into suspended links. In a previous study [2], when a loop occurred on an Ethernet network, incorrect learning of the MAC address table on the Ethernet switch was recovered and a loop point was detected. Moreover, link aggregation [3] was proposed for effective utilization of physical links when there is loop topology on a network. The link aggregation technology is used when there are plural physical links between neighboring Ethernet switches. This link aggregation technology combines these plural physical links into one logical link and then does not suspend the links like the STP. A control technology that limits an Ethernet frame round to a finite amount of time has also been proposed [4]. This technology appends a time-to-live (TTL) field to each Ethernet header on an overlay Ethernet (EoE) network.

A. STP

The STP constructs a virtual tree network [Fig. 1 (b)] on an Ethernet that has loop topology [Fig. 1 (a)] by exchanging the bridge protocol data units (BPDUs) [Fig. 2] between neighboring Ethernet switches. These BPDUs change the Ethernet frame round.

<table>
<thead>
<tr>
<th>Ports mode</th>
<th>Behavior</th>
<th>Receive BPDU</th>
<th>Send BPDU</th>
<th>Learn MAC Address</th>
<th>Transport Ethernet Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disabled</td>
<td>Blocking</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Listening</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Learning</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Transport</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
switch port modes to one of five modes as shown in Table I to construct a virtual tree network on the Ethernet.

When a virtual tree network is first being constructed with the STP, all switch ports are in blocking mode except for the disabled mode ports, which are configured by an operator. Then, some of these change to listening mode when the blocking mode switch ports do not receive BPDUs within a specific amount of time. When the switch port is in the listening mode, it can send and receive BPDUs to and from neighboring switches. Then, the highest priority (lowest bridge ID) Ethernet switch is selected as a root bridge. After the root bridge is selected, each Ethernet Switch chooses the lowest cost (root path cost) port as the route to the root bridge. Each Ethernet switch port that is chosen as a lowest cost route is changed to transport mode after learning mode. With this STP protocol, a virtual tree network is constructed. With this virtual tree network construction process, some physical links are suspended. However, one of the Ethernet switch ports that has both ends of a suspended link is not changed to blocking mode. At both end ports, a port of a lower root path cost is still in listening mode, and another port is changed to blocking mode. After the virtual tree network is constructed, neighboring switches periodically exchange BPDUs. When network topology is modified by network trouble or the addition of a new node, the STP reconstructs a virtual tree network using every physical link, which include suspended links.

### Table II: The relation between behavior mode and function.

<table>
<thead>
<tr>
<th>Bypass link status on STP</th>
<th>Ethernet NAT</th>
<th>Bridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bypass mode</td>
<td>Enable</td>
<td>Disable</td>
</tr>
<tr>
<td>Pass-through mode</td>
<td>Disable</td>
<td>Enable</td>
</tr>
</tbody>
</table>

### B. PBB-TE

The provider backbone bridge traffic engineering (PBB-TE) [4] technology controls Ethernet frame routes on a special EoE network. The PBB-TE network is constructed by a core network and several edge Ethernet switches. Each edge Ethernet switch is placed at the border between the core network and the general Ethernet network. The core network configures multiple paths between edge Ethernet switches using a virtual LAN (VLAN). An ingress edge Ethernet switch encapsulates the Ethernet frame that is received from the general Ethernet network, and the edge Ethernet switch assigns a suitable path by appending a VLAN ID to the encapsulated Ethernet frame. The encapsulated Ethernet frame is decapsulated by an egress edge Ethernet switch when the frame leaves the core network. Then, the decapsulated Ethernet frame is transported to the general Ethernet network. As mentioned, the PBB-TE can create and use multiple paths on the Ethernet. However, tagged-VLAN is required for every core network and edge switch to operate PBB-TE.

### C. Link aggregation

Link aggregation (IEEE 802.3ad, IEEE 802.1ax) is used for enabling more than a single physical link’s bandwidth by combining plural physical links into one logical link. However, these plural physical links, which are components of the logical link, must be connected directly between Ethernet switches that operate the link aggregation. Moreover, the link aggregation cannot be operated when the loop topology is constructed by some Ethernet switches that include no link aggregation Ethernet switch.

### D. Problem of Ethernet routing

The STP prevents Ethernet frame round on the Ethernet. Link aggregation and PBB-TE help effectively operate loop topology physical links by not suspending links like the STP. However, communication between Node a and Node c pass through a roundabout route when Link Z is suspended by STP, as shown in Fig. 1 (b). Then, even if Link Z has enough bandwidth for communication between Node a and Node c, communication between Node a and Node c cannot take enough bandwidth when Ethernet Switch B is congested, and the link aggregation requires physical links that are connected directly between link aggregation Ethernet switches. The PBB-TE requires that all Ethernet switches in a PBB-TE network have PBB-TE function. Moreover, the Ethernet switches that implement these functions are generally expensive. Therefore, applying these high-functional Ethernet switches to an existing network is not easy.

### III. Proposed method

We describe an Ethernet Bypass Node (EBN) to solve this problem. The EBN is for Ethernet networks constructed with low-function Ethernet switches that do not have tagged-VLAN functions. The EBN is placed on the suspended link by STP.
The relation between the EBN state and network state.

in usual which the operator hopes to use as a bypass. The suspended links used to construct a virtual tree network with the STP are used for backup if there is network trouble. Then, even if the EBN is placed at a suspended link and the suspended link is used for an Ethernet bypass route, the suspended link should be used as a backup link by STP if there is trouble. The EBN has two behavior modes as shown in the following list. The EBN behavior mode changes depending on the situation.

- Bypass mode
- Pass-through mode

These bypass and pass-through modes are changed by enabling or disabling the Ethernet Network Address Translation (NAT) function and bridge function, as shown in Table II. Here, the Ethernet NAT is the function of modifying the MAC address of the Ethernet frame. Section III-B2 explains the Ethernet NAT behavior on the EBN. The bypass mode and pass-through mode behavior and mode change mechanisms are presented in Sections III-A, III-B and III-C.

A. Mode change mechanism

The EBN should guess the network condition by observing the network information. Therefore, the EBN observes Ethernet frames that are received from both neighboring ports connected to the EBN. The EBN counts the number of BPDUs and other general Ethernet frames received from each port to decide the EBN behavior mode. The relation among the EBN behavior modes, network states, and the number of BPDUs and other general Ethernet frames received from each port is shown in Fig. 3. Fig. 4 is a state transition diagram of the states shown in Fig. 3.

In Fig. 4, when the EBN is appended at a link on the Ethernet network operated by STP, the EBN mode starts from d) pass-through mode and is then transmitted to another mode depending on the network condition.

Then, when the EBN is placed in advance at a link on an Ethernet network operated by STP, the EBN mode starts from a) bypass mode and then changes immediately to e). After that, the EBN mode is changed to d) pass-through mode when a virtual tree network starts to be constructed by STP.

When the EBN is placed in advance at a link on an Ethernet network that is not operated by STP, the EBN mode starts from a) bypass mode and then changes immediately to e). After that, the EBN mode does not change from e) because the network state does not change.

The EBN behavior mode is decided as described.

B. Bypass mode

The EBN has a table of bypass targets. The table of bypass targets records pairs of IP address for bypass target nodes and...
forwarding ports for the bypass target nodes in advance. Then, in bypass mode, Ethernet bypass is activated by the following two functions, in accordance with the table of bypass targets.

- Control of the bypass target node’s ARP table.
- Ethernet network address translation.

Next, we will explain the bypass mode behavior using an actual example as shown in Fig. 5.

1) ARP table control for bypass target nodes: As shown in Fig. 6, a suspended link is not used for communication between Node a and Node c on a virtual tree network constructed by STP. Therefore, when the EBN is placed at a link that the operator hopes to use as a bypass as shown in Fig. 5, communication between Node a and Node c should pass not through the virtual tree route determined by STP but through the Ethernet bypass link determined by EBN. Ethernet frames communicated between Node a and Node c should be transported to the bypass ports connected to the EBN on Ethernet Switch A and Ethernet Switch B. Therefore, for activating the bypass communication between Node a and Node c, the EBN overwrites the ARP table of EBN Node a and Node c as the bypass target nodes.

The EBN’s ARP table overwrite behavior is as follows. An actual example is shown in Fig. 5. On an Ethernet network as shown in Fig. 5, when Node a sends an IP packet to I[c] (IP address of Node c), Node a sends an ARP request to the entire network to resolve to M[c] (MAC address of Node c) from I[c] [Fig. 5(1)]. Then, when Node c receives the ARP request from Node a, Node c sends “I[c] is at M[c]” as an ARP reply to Node a [Fig. 5(2)]. When Node a receives the ARP reply, Node a writes “I[c] is at M[c]” as the ARP reply to Node a [Fig. 5(2’)]. If there is trouble with a link on a virtual tree network, there is a possibility that the link placed the EBN will be used for a newly reconstructed virtual tree network by STP. Then, if the EBN only has bypass mode, the link cannot be used for virtual tree network reconstruction because the bypass mode of the EBN does not transport the BPDUs. The exchange of BPDUs between Ethernet switches is necessary for virtual tree network reconstruction. Therefore, the EBN should transport not only BPDUs but also all Ethernet frames connected to EBN p1, and are not transported to p2 side, which is the route determined by STP.

2) Ethernet NAT for IP packets: As presented above, when a network configuration is shut down as in Fig. 5, the EBN changes the destination Ethernet frame address from Fig. 7(a) to Fig. 7(b) by overwriting the ARP table of bypass target nodes. Then, IP packets sent from Node a to Node c are output to p3 of Ethernet Switch A, which is connected to EBN. If EBN p1 receives these packets from Node a and then outputs the received packets to EBN p2, the output packets do not reach Node c but return to EBN p1 via Ethernet Switch C, Ethernet Switch B, and Ethernet Switch A, because the destination MAC address of the packets output by EBN p2 is M[Ep1]. Therefore, when the EBN receives a packet addressed to Node c from Node a as communication from a bypass target, the EBN translates the destination MAC address from M[c] to M[Ep1] and also translates the source MAC address from M[a] to M[Ep2] to deliver the packet for Node c to Node c correctly, and then outputs the packets to EBN p2. This output packet from EBN 2 is delivered to Node c via Ethernet Switch C.

In this way, the EBN can bypass only the communication between selected nodes.

The behavior of the EBN disperses traffic on Ethernet networks constructed by tagged-VLAN disabled Ethernet switches.

C. Pass-through mode

The EBN pretends just a simple link as pass-through mode when the STP constructs a virtual tree network or the EBN is appended to a link used for the STP’s virtual tree network as presented in III-A.

The link that the EBN places is considered to be a suspended link by STP as usual. However, when there is trouble with a link on a virtual tree network, there is a possibility that the link placed the EBN will be used for a newly reconstructed virtual tree network by STP. Then, if the EBN only has bypass mode, the link cannot be used for virtual tree network reconstruction because the bypass mode of the EBN does not transport the BPDUs. The exchange of BPDUs between Ethernet switches is necessary for virtual tree network reconstruction. Therefore, the EBN should transport not only BPDUs but also all Ethernet frames connected to EBN p1, and are not transported to p2 side, which is the route determined by STP.

Fig. 5: Communication on network constructed by STP.

(a): Normal Ethernet frame.

<table>
<thead>
<tr>
<th>MAC</th>
<th>MAC</th>
<th>IP</th>
<th>IP</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>a</td>
<td>a</td>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>

(b): Ethernet frame from overwritten ARP Table Node a.

<table>
<thead>
<tr>
<th>MAC</th>
<th>MAC</th>
<th>IP</th>
<th>IP</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBN p1</td>
<td>a</td>
<td>a</td>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>

(c): Ethernet frame from EBN NAT.

<table>
<thead>
<tr>
<th>MAC</th>
<th>MAC</th>
<th>IP</th>
<th>IP</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>EBN p2</td>
<td>a</td>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5: Communication on network constructed by STP.
frames when the EBN transports BPDUs, and the links placed by the EBN are used as part of the virtual tree network by STP.

Therefore, the EBN should be implemented not only in bypass mode but also in pass-through mode which transports all Ethernet frame.

IV. EXPERIMENT

The effect and the behavior of the EBN on the actual network were tested by experiment as shown in Fig. 8. We generate a bottleneck link between Ethernet Switch A and Ethernet Switch B by generating background traffic between Node m and Node n as shown in Fig. 8. Then, we observed packet drop between Node a and Node b while changing the background traffic when using a conventional virtual tree route without the EBN and using bypass route with the EBN.

The results showed that when using a conventional route selected by STP with a bottleneck, packet drop occurred between Node a and Node b due to background traffic growth, shown as “Without EBN” in Fig. 9. On the other hand, when using a bypass route that used suspended links with the EBN, packet drop did not occur between Node a and Node b and the communications were always stable despite background traffic growth, shown as “With EBN” in Fig. 9.

V. DISCUSSION

In this section, we will discuss the change mechanism of the Ethernet frame route on the Ethernet switch.

As a technique for transporting Ethernet frames between Node a and Node b to the EBN side on Ethernet Switch A and Ethernet Switch B in Fig. 5, we may use the MAC address auto-learning function on an Ethernet switch instead of overwriting the ARP table of bypass target nodes with the EBN. For example, in Fig. 5, EBN p1 may spoof the source MAC address of an Ethernet frame (change to M[c]), and then Ethernet Switch A, which receives the spoofed frame, learns that the Ethernet frame that has the destination address Node c is sent to p3 which is connected to the EBN side. However, even though the next port of the Ethernet frame that has the destination Node c on Ethernet Switch A is changed by the above technique, the MAC address table on Ethernet Switch A, which is learned as explained, is overwritten again. Therefore, for example, Node c sends a broadcast frame and then the Ethernet Switch A receives the broadcast frame that has the source MAC address M[c]. Of course, it is possible to solve the problem that EBN p1 may spoof the source MAC address of the Ethernet frame (changed to M[c]) again whenever the learning result is overwritten. However, there is a possibility that this overwriting learning result by broadcast will happen frequently and then the communication quality may become unstable because the Ethernet broadcast frame is frequently broadcast for ARP requests and so on.

Therefore, in this research, the EBN overwrote the ARP table of the bypass target node and did not change the MAC table on the Ethernet switch to change the Ethernet virtual tree route.

VI. CONCLUSION

We described an Ethernet Bypass Node (EBN) for traffic dispersion. The EBN creates an Ethernet bypass, and then only selected communications between specific nodes can pass through link suspended by STP. Moreover, for the STP’s virtual tree network reconstruction, the EBN provides two modes, bypass mode and pass-through mode. These modes are changed depending on the network state.

Then, we implemented the EBN on a PC and experimented on an actual network to test the behavior and effectiveness of the EBN. Thus, we showed that the EBN improves communication quality by using links suspended by STP.

In the future, bypass target node selection and the EBN configuration and timing of sending ARP replies from the EBN should be automated.

ACKNOWLEDGMENTS

This work was partly supported by the Global COE Program, "Photonics Integration-Core Electronics", MEXT, Japan.

REFERENCES