INTERNET-DRAFT Expires: September 2004

T. Ogura M. Maruyama NTT Network Innovation Labs T. Yoshida Werk Mikro Systems March 2004

A Multicast Extension to MAPOS NSP (Node Switch Protocol)

<draft-ogura-mapos-nsp-multiexp-03.txt>

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/1id-abstracts.html

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

Distribution of this memo is unlimited. Please send comments to the authors <ogura@core.ecl.net> <mitsuru@core.ecl.net> and <yoshida@peta.arch.ecl.net>.

Abstract

Multiple Access Protocol over SONET/SDH (MAPOS) is a link-layer protocol for transmitting network-layer datagrams encapsulated in High-Level Data Link Control (HDLC) frames over Synchronous Optical NETwork/Synchronous Digital Hierarchy (SONET/SDH).

This document describes an extension to the Node Switch Protocol (NSP) of MAPOS. NSP is a protocol for automatically assigning addresses to MAPOS network interfaces of nodes. The extension NSP+ enables the nodes to provide their directly connected switches with information about the multicast groups to which the MAPOS network

Ogura, et al. Expires September 2004

interfaces of the nodes belong, so a switch forwards only required multicast frames to the network interfaces.

1. Introduction

Multiple Access Protocol over SONET/SDH (MAPOS) [1][2] is a link-layer protocol for transmitting High-Level Data Link Control (HDLC) frames over SONET/SDH [6][7]. In MAPOS, HDLC frames (MAPOS frames) are switched according to their destination addresses and eventually delivered to the destination network interfaces of nodes. (The term "nodes" means components other than switches in MAPOS networks; nodes are usually hosts or IP routers.)

In addition to unicasting, MAPOS also supports multicasting. Namely, the destination address field of an HDLC frame can include a multicast address which indicates that the frame must be delivered to one or more network interfaces of nodes. Conventionally, even when an optimized multicast transmission tree had been built in the network layer (layer 3), MAPOS switches forwarded multicast frames to all ports except for receiving ports regardless of their destination addresses because there were no specifications to restrict unnecessary multicast-frame transmission in MAPOS. However, it is preferable for the switches to forward multicast frames only to ports to which the frames really need to be forwarded, instead of flooding all ports with them.

This document describes NSP+, which eliminates unnecessary multicast-frame transmission between a MAPOS switch and its directly connected nodes. Similar to NSP [3], NSP+ is a protocol which enables addresses to be automatically assigned to each MAPOS network interface of nodes. In addition, NSP+ enables nodes to provide their directly connected switches with information about the multicast groups to which the MAPOS network interfaces of the nodes belong, so a switch forwards only required multicast frames to the network interfaces.

In the remainder of this document, the term "MAPOS" is used unless the distinction between MAPOS version 1 $[\underline{1}]$ and MAPOS 16 $[\underline{2}]$ is required.

2. NSP+

2.1 Overview

Similar to NSP, NSP+ provides an automatic network-interface address assignment function in MAPOS. It reduces the administrative burden of network-interface address configuration and prevents trouble such as address inconsistency and collision. When a MAPOS network interface of a node is connected to a switch and receives a SONET signal

[Page 2]

correctly, the node sends an address request frame from the network interface to the control processor in the switch; the destination address of this frame is 0x01 (MAPOS version 1) [1] or 0x0001 (MAPOS 16) [2].

NSP+ differs from NSP in that an optional multicast field can be included in the address request frame. The multicast field contains a complete set of multicast MAPOS addresses derived from all the addresses of the network-layer multicast groups to which the network interface sending the address request frame belongs. (Device controller or operating system software usually holds such a set of link-layer multicast addresses on each network interface.)

When the switch receives the address request frame, it replies with an address assignment frame. The address assignment frame is exactly the same as that of NSP, i.e., it includes the assigned MAPOS address, and its destination is the assigned address [3]. The address to be assigned to the network interface is decided in the same way as in NSP, i.e., the address consists of the switch number and the port number of the switch to which the network interface is connected [3].

In addition, the switch configures its frame forwarding table so that unicast MAPOS frames destined to the address that has just been assigned to the network interface are transmitted to it, in the same way as in NSP.

Differently from NSP, in NSP+ the switch also configures the frame forwarding table so that in the case of multicasting, it forwards to the network interface only multicast frames whose destination addresses are the same as one of the multicast MAPOS addresses included in the multicast field of the address request. This eliminates unnecessary multicast-frame transmission between a switch and its directly connected nodes.

NSP+ does not influence the transmission of broadcast frames or the frame transmission between switches.

2.2 Frame format

The HDLC protocol field of an NSP+ frame contains 0xFE03 (hexadecimal) similar to NSP [5], and the HDLC information field contains a 32-bit command field and a 32-bit address field shown in Figure 1, which are the same as those of NSP [3].

[Page 3]

+----+ | Command | Address | +----+ |<- 32 bits ->|<- 32 bits ->|

Figure 1. Content of the HDLC information field of an NSP+ frame.

The details of the fields are as follows:

Command	1 -	address request
	2 -	address assignment
	3 -	reject (error)

Address In address request frames, the address field should be filled with zeros, although the switch ignores it. In address assignment frames, in the case of MAPOS version 1, the assigned 8-bit address is placed in the least significant octet of the field and the rest of the field is filled with zeros. In the case of MAPOS 16, the assigned 16-bit address is placed in the least significant octets of the field and the rest of the field is filled with zeros. if the switch cannot assign the address for some reason, the switch replies with a reject frame (the value of the command field is 3). The value of its address field is undefined.

In NSP+, differently from NSP, when the value of the command field is 1 (address request) and the address field is filled with zeros, an optional multicast field can be added as shown in Figure 2.

+----+ | Command | Address | Multicast | +-----+ |<- 32 bits ->|<- 32 bits ->|

Figure 2. Content of the HDLC information field of an NSP+ frame with a multicast field.

The multicast field contains multicast MAPOS addresses derived from all the addresses of the network-layer multicast groups to which the MAPOS network interface sending this address request frame belongs. The length of this field is variable, and only one multicast field can be added regardless of the number of included multicast MAPOS addresses. This field is optional; an address request frame that has

[Page 4]

no multicast field can also be used (see (b.3) in section 2.3).

Figure 3 shows the details of the multicast field. The address fields store multicast MAPOS addresses derived from addresses of the network-layer multicast groups to which the network interface belongs.

Figure 3. Multicast field format.

The details of the fields are as follows:

Code	1 - reserved for future use 2 - multicast
	(Indicates that the multicast field includes multicast MAPOS addresses.) Other values are not used.
Form	1 - MAPOS version 1

2 - MAPOS 16

Length The length of the multicast field in octets.

Address Includes a multicast MAPOS address. In MAPOS version 1, an 8-bit multicast MAPOS address is placed in the least significant octet of the 32-bit address field and the rest of the field is filled with zeros. In MAPOS 16, a 16-bit multicast MAPOS address is placed in the least significant octets of the 32-bit address field and the rest of the field is filled with zeros.

There is no limit on the number of address fields in a multicast field, as long as the length of the content shown in Figure 2 does not exceed the size of the maximum transmission unit (MTU) of the network (see (b.2) in section 2.3). In addition, the number of address fields can be zero (see (b.4) in section 2.3).

2.3 Behavior of nodes

A node must configure multicast fields and send address request frames as described below.

[Page 5]

(a) Generating multicast MAPOS addresses

Each multicast MAPOS address included in a multicast field is generated from a network-layer multicast group address in the way that each "<network layer> over MAPOS" specification describes. For example, in the case of IPv4 over MAPOS version 1 [4], the MSB and LSB of the MAPOS address are set to 1, and the remaining six bits of the MAPOS address must contain the lowest-order six bits of the IP multicast group address [8]. (Here, the term "multicast" does not include broadcast. A multicast field must not include the broadcast MAPOS address.)

As in the case of IP, multicast group addresses can be of two types. The first is addresses corresponding to IP multicast groups to which all the network interfaces of nodes have belonged since the time of system initialization, such as 224.0.0.1 (all hosts) or 224.0.0.2 (all routers) in IPv4. The other is group addresses which a network interface joins or leaves at the request of application software. Regardless of that, in order for the nodes to receive network-layer datagrams destined to these addresses, corresponding multicast MAPOS addresses must be included in the multicast field.

- (b) Configuring a multicast field
- (b.1) The set of destination addresses of multicast MAPOS frames that a node can receive through a MAPOS network interface is exactly the same as the set of multicast MAPOS addresses included in the multicast field of the latest address request frame sent from the network interface. (If the latest address request frame does not have a multicast field, the node can receive all multicast frames regardless of their destination addresses through the network interface. See (b.3).)

Therefore, whenever a network interface sends an address request frame that has a multicast field, the multicast field must contain a complete set of multicast MAPOS addresses derived from all the addresses of the network-layer multicast groups to which the network interface belongs. The network interface can start or stop receiving multicast MAPOS frames having certain multicast MAPOS addresses by sending a new address request frame that has a new complete set of multicast MAPOS addresses in its multicast field.

(b.2) There is no limit on the number of multicast MAPOS addresses included in a multicast field. In other words, all the multicast MAPOS addresses in the MAPOS address space can be included in a single multicast field. This implies that the largest possible multicast field contains 64 (the sixth power

[Page 6]

of two) addresses in the case of MAPOS version 1 and 8192 (the thirteenth power of two) addresses in the case of MAPOS 16.

In the latter case, the address fields of the multicast field amounts to 32,768 octets in total because each multicast MAPOS address is stored in a 32-bit-wide field, as described in <u>section 2.2</u>. Even in this case, the content shown in Figure 2 can be accommodated by an address request frame, because the maximum length of the information field of a MAPOS frame is 65,280 octets [1][2].

However, in a MAPOS implementation where the size of the maximum transmission unit (MTU), i.e., the maximum length of the information field of a MAPOS frame, is smaller than the length of the content shown in Figure 2 with its largest possible multicast field, it is possible for the length of the content to exceed the size of the MTU. Such content cannot be transmitted because NSP+ does not have a function for dividing it into multiple address request frames.

In such cases, a node must send an address request frame that has no multicast field through the network interface in order to make the directly connected switch forward all multicast frames to it regardless of their destination addresses (see (b.3)). Even though this causes some unnecessary multicast frame transmission, the node can receive all the multicast frames it really needs to receive through the network interface.

(b.3) When a node needs to receive all multicast frames regardless of their destination addresses through a network interface, it must send an address request frame that has no multicast field through the network interface, i.e., an address request frame that only contains the fields shown in Figure 1.

For example, this rule defines the outcome of cases where a node that supports only conventional NSP is connected to a switch that supports NSP+. According to this rule, the node receives all multicast frames regardless of their destination addresses.

(b.4) When a node does not need to receive any multicast frames through a network interface, it must send an address request frame including a multicast field that has no address fields through the network interface, i.e., a multicast field consisting of code, form, and length field. In this case, the value of the length field is 4.

[Page 7]

In the case of IPv4, this is an exceptional case because a node's network interface which supports IP multicasting usually belongs to multicast group 224.0.0.1 (all hosts) or 224.0.0.2 (all routers).

(c) When to send an address request frame

In NSP+, address request frames are sent basically at the same events as those for NSP. These are described in [3] as follows.

- (c.1) When a network interface of a node is connected to a switch and receives a SONET signal, it sends an address request frame to the switch. When the switch receives the frame, it replies with an address assignment frame.
- (c.2) If the network interface has not yet been assigned a MAPOS address and it does not receive the address assignment frame within 5 seconds, it continues to send address request frames until it receives the address assignment frame.
- (c.3) Whenever a network interface of a node detects a transmission error such as carrier loss or out-of-synchronization, it should send an address request frame to the switch and verify its current address.
- (c.4) In addition to (c.3), a network interface must verify its address by sending address request frames every 30 seconds. The switch regards these as keep-alive frames and utilizes them to detect the interface's status. If the switch does not receive an address request frame for more than 90 seconds, it assumes that the network interface went down.

Moreover, in NSP+, different from NSP, an address request frame should be sent under the following condition.

(c.5) When a network interface joins or leaves network-layer multicast groups at the request of application software running on a node, an address request frame should be sent from the network interface. Here, its multicast field must contain the new complete set of multicast MAPOS addresses derived from all the addresses of the new network-layer multicast groups to which the network interface belongs.

2.4 Behavior of switches

A switch compliant with NSP+ must obey the following when it receives an address request frame from a network interface of a directly connected node:

[Page 8]

- 1) A switch decides a MAPOS address to be assigned to the network interface in the same way as in NSP, then replies with an address assignment frame. This address assignment frame is exactly the same as that of NSP. If the switch cannot assign the address for some reason, it replies with a reject frame.
- 2) A switch must configure its frame forwarding table so that unicast MAPOS frames destined to the address that has just been assigned to the network interface are transmitted to it and so that multicast MAPOS frames are transmitted to the network interface as described in (b.1), (b.3), and (b.4) in <u>section</u> 2.3.

The configuration of the frame forwarding table relevant to the transmission of broadcast frames or to the frame transmission between switches is not controlled by NSP+.

The order of the above actions is not specified. Furthermore, for a switch that supports only NSP, its behavior upon receiving an address request frame that includes a multicast field is not defined.

<u>2.5</u> Consideration for special cases

There are two special cases to consider as described in [3]. The first is point-to-point connection of two nodes without a switch. The other is loop-back, that is, direct connection between the input and the output of the same port of a network interface of a node.

In these cases, each network interface is assigned an address 0x03 (MAPOS version 1) or 0x0003 (MAPOS 16) in the same way as in NSP. The multicast field of an address request frame is ignored.

3. Example

Figure 4 shows an example. In Figure 4, IPv4 over MAPOS version 1 is assumed and the switch numbers are assumed to be two bits long. Node N1 is connected to port 0x3 of switch S1 and node N2 is connected to port 0x5 of the same switch. The switch is numbered 0x1 (01 in binary).

Ogura, et al. Expires September 2004 [Page 9]



Figure 4. An example of NSP+.

The network interface of node N1 is a member of IPv4 multicast groups G1 and G2, and that of N2 is a member of IPv4 multicast groups G1 and G3. Here, G1 stands for an IPv4 multicast group address and G1' stands for a multicast MAPOS address derived from G1; the same applies to the others. In addition, the notation "m_field (G1', G2')" stands for a multicast field which includes multicast MAPOS addresses G1' and G2'.

After joining their multicast groups, N1 and N2 send address request frames (AddrReq) with multicast fields to the control processor in S1 through their network interfaces. Specifically, N1 and N2 send address request frames with m_field (G1', G2') and m_field (G1', G3'), respectively, and the destination of the address request frames is 0x01 (the control processor in the local switch).

On receiving the address request frames, S1 decides the MAPOS addresses to be assigned to the network interfaces and sends back address assignment frames (AddrAssign) in the same way as NSP does. The assigned addresses are of the form "O <switch number> <port number> 1" [3]. In this example, the address assigned to N1 is O + O1 + OOO11, that is, OO1OOO11 (Ox23), and the address assigned to N2 is OO1OO1O1 (Ox25). The address assignment frames include these addresses, and their destinations are Ox23 and Ox25, respectively.

Ogura, et al. Expires September 2004 [Page 10]

In addition, S1 configures its frame forwarding table so that it forwards incoming unicast frames with destination address 0x23 to N1 and frames with destination address 0x25 to N2 respectively, and so that it forwards incoming multicast frames with destination address G1' to N1 and N2, frames with destination address G2' only to N1, and frames with destination address G3' only to N2. It also ensures that any multicast frames with destination addresses other than G1', G2', and G3' are not forwarded to N1 or N2.

Care should be taken that NSP+ does not influence frame transmission between switches. Usually in MAPOS, all multicast frames are transmitted between S1 and S2.

<u>4</u>. Security Considerations

This section describes security factors related to NSP $[\underline{3}]$ and NSP+. Common factors of NSP and NSP+ are discussed in <u>section 4.1</u>, and an NSP+ specific factor is discussed in <u>section 4.2</u>.

4.1 Common factors of NSP and NSP+

4.1.1 Correspondence of an interface's MAPOS address to its location

In a multi-switch environment, a MAPOS address of a network interface of a node assigned via NSP or NSP+ consists of the switch number and the port number of the switch to which the network interface is connected. This brings the following advantages.

- 1. The value of the MAPOS address of a network interface of a node indicates the location of the interface in the MAPOS network. In other words, the destination address of a unicast MAPOS frame defines the actual location to which the frame should finally be delivered. Therefore, as long as MAPOS addresses of network interfaces of nodes that have been connected to the network through proper administrative process are held, and as long as it is ensured that in the case of unicasting only frames destined to those addresses are delivered (e.g., by using destination address filtering), other nodes cannot receive frames unless their network interfaces are connected to the same ports of switches as those to which network interfaces of properly administered nodes are connected. This makes fraudulent reception of unicast traffic difficult.
- 2. In the case where MAPOS addresses are not administered as mentioned above, a malicious node could hijack traffic by spoofing its layer-3 address in a response to a link-layer address resolution. Even in this case, the node must advertise the true MAPOS address of its network interface in the response so that it

Ogura, et al. Expires September 2004 [Page 11]

can receive successive frames. This makes it easy to pinpoint the location of the node.

4.1.2 A flood of address requests

An address request frame should be sent from nodes only when one of the events described in <u>section 2.3</u>(c) occurs. However, a malicious or malfunctioning node could send a large number of address request frames in a very short time to the control processor of its directly connected switch. In this case, if the receiving switch has only straightforward functions to support NSP or NSP+, the load on its control processor for processing these address request frames would become high and this might lead to system failure.

In order to prevent the above, a switch should have a function to monitor the arrival frequency of address request frames from each port to its control processor, and to disconnect a line if a node connected to it has sent too many address request frames.

4.2 An NSP+ specific factor

4.2.1 Insertion of unicast addresses in a multicast field

In each address field of a multicast field of an address request frame, a multicast MAPOS address must be inserted, but a malicious or malfunctioning node could insert a unicast MAPOS address in it. In this case, if the receiving switch has only a straightforward function to configure its frame forwarding table so that it forwards MAPOS frames destined to the unicast address to the port from which the address request frame arrived, the unicast MAPOS frames will consequently be sent to the port, which is not necessarily their real destination port. This leads to inappropriate or fraudulent reception of unicast traffic.

In order to prevent the above, a switch should have a function to remove or ignore MAPOS addresses which are not multicast addresses included in a multicast field of an address request.

5. Normative References

- [1] Murakami, K. and M. Maruyama, "MAPOS Multiple Access Protocol over SONET/SDH, Version 1," <u>RFC-2171</u>, June 1997.
- [2] Murakami, K. and M. Maruyama, "MAPOS 16 Multiple Access Protocol over SONET/SDH with 16 Bit Addressing," <u>RFC2175</u>, June 1997.
- [3] Maruyama, M. and K. Murakami, "A MAPOS Version 1 Extension -

Ogura, et al. Expires September 2004 [Page 12]

Node Switch Protocol," <u>RFC-2173</u>, June 1997.

- [4] Murakami, K. and M. Maruyama, "IPv4 over MAPOS Version 1," <u>RFC-2176</u>, June 1997.
- [5] Murakami, K. and M. Maruyama, "MAPOS Version 1 Assigned Numbers," <u>RFC-2172</u>, June 1997.

<u>6</u>. Informative References

- [6] American National Standards Institute, "Synchronous Optical Network (SONET) - Basic Description Including Multiplex Structure, Rates and Formats," ANSI T1.105-1995.
- [7] ITU-T Recommendation G.707, "Network Node Interface for the Synchronous Digital Hierarchy (SDH)," October 2000.
- [8] Deering, S., "Host Extensions for IP Multicasting," <u>RFC-1112</u>, August 1989.

7. Authors' Addresses

Tsuyoshi Ogura NTT Network Innovation Laboratories 3-9-11, Midori-cho Musashino-shi Tokyo 180-8585, Japan E-mail: ogura@core.ecl.net

Mitsuru Maruyama NTT Network Innovation Laboratories 3-9-11, Midori-cho Musashino-shi Tokyo 180-8585, Japan E-mail: mitsuru@core.ecl.net

Toshiaki Yoshida Werk Mikro Systems 250-1, Mikajiri Kumagaya Saitama 360-0843, Japan E-mail: yoshida@peta.arch.ecl.net

8. Full Copyright Statement

Copyright (C) The Internet Society (2004). All Rights Reserved.

This document and translations of it may be copied and furnished to

Ogura, et al. Expires September 2004 [Page 13]

others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Ogura, et al. Expires September 2004 [Page 14]