Multicasting for Token Ring/Ethernet

Additional IGMP/IP support

Mon, Jul 13, 1998

Multicasting support is required for communication between networks via IP. The Internet protocol standard that addresses this need is IGMP, the support for which has been included for IRMs for a long time. This has meant that nodes that connect to different physical ethernets can communicate using IP multicast addressing. This capability means that a node can query the network for the location of a given device name, or it can send a request for information that spans multiple nodes and have that request message delivered to all nodes across multiple networks, or it can deliver an alarm message that can be received by nodes across multiple networks. It requires that IP multicasting support exist in the router software.

Let's review how IGMP accomplishes its objectives. Every minute, the router sends out an IGMP query message to each of its connected networks in order to learn what multicast IP addresses are of interest to nodes on each network. (A multicast router only forwards multicast datagrams to those networks for which at least one node has indicated an interest.) Each node on a network that supports IGMP listens to the all-hosts multicast destination IP address, which is 224.0.0.1. In order that the network is not swamped with a plethora of replies, each node schedules a reply message (for each of the IP multicast addresses for which it has an interest) to be sent after a random delay over the next 10 seconds. The reply message is sent to the same IP multicast destination. If a node receives such a reply (that was transmitted by another node on its own network), it cancels its intention to send any pending reply for the same IP multicast address. In this way, the reply network activity is kept to a minimum. The router does not need to know which nodes on a network have an interest in a given multicast address; it only needs to know if *any* node has such an interest.

All of the above support has existed for ethernet-based nodes for some time, but it had not been implemented for token ring nodes. Recently, such token ring multicasting support has been provided with an upgrade of the router software, thus making it useful to support IGMP for token ring as well as ethernet. The aim is to allow an ethernet node to look up a device name, say, that resides in a token ring node. It also allows a token ring node to look up a device name on a different token ring that is only reachable (via IP) through a router. Heretofore, the connectivity between different token ring networks has been accomplished by bridging support only. Once the router support was enabled for IP multicasting, this support was inadequate to the task.

IP multicast support for the token ring nodes is implemented, insofar as is reasonably possible, in the same way as was done for ethernet. A table of up to 8 *ethernet*-mapped multicast physical addresses is maintained in system table TRING at offset `0240`. A simple example of the contents of this table is as follows:

```
105240     0100  5E00  0001  0013
105248     0100  5E00  02F9  0027
105250     0000  0000  0000  0000
105258     0000  0000  0000  0000
105260     0000  0000  0000  0000
```

```
105268     0000  0000  0000  0000
105270     0000  0000  0000  0000
105278     0000  0000  0000  0000
```

Each entry is 8 bytes in length, the first 6 of which are used to house the ethernet-mapped hardware address. The rule for building such an address is to append the low 23 bits of the IP multicast address to high 25 bits of the fixed ethernet multicast physical address `0100 5E00 0000`. For the token ring support, this same scheme is used for the entries in this table, in order to maintain compatibility with existing software. But for token ring IP multicasting, only a single 6-byte hardware address is actually used in the hardware, `C000 0004 0000`. Any IP multicast address is mapped to this fixed group functional token ring destination address for transmission.

The last two bytes of each entry are used for counters. The high byte of the last word is used as a delay counter, set to a random count in the range 1–100 when the router's IGMP query message is received. Each 10–15 Hz cycle, any nonzero count is decremented; if the count reaches zero, a reply message is queued to the network. If an IGMP reply message is received from another node for a given IP multicast address—implying that the other node chose a smaller random delay count—the corresponding delay counter is cleared, thereby cancelling the intended reply. The low byte of the last word is a diagnostic counter that is incremented when a datagram is received addressed to the corresponding IP multicast address.

The previous table is used to filter incoming datagrams. If a datagram is received that targets an IP address in the multicast range of 224.0.0.1—239.255.255.255, a check is made against the table to determine whether it should be accepted. This is necessary since all IP multicasting on token ring shares the same physical address of `C000 0004 0000`. Unlike the case of ethernet, in which the hardware can do at least a major part of the filtering, the token ring case requires software filtering of all IP multicast received transmissions.

In order to receive IP multicast datagrams, the token ring controller must be enabled to receive frames addressed to the fixed group functional address that is used. Again, in the TRING table, the specification for which group functional addresses can be used is contained in the 32-bit long word at `0010504C`. To enable hardware acceptance of the group functional address `C000 0004 0000`, bit 18 of this long word must be set. In Linac node `0614`, for example, we find a value of `44040000` for this long word. (The other two bits set allow acceptance of `C000 4000 0000` and `C000 0400 0000`, which are used for certain non-IP network processing by Linac nodes.)

The table that specifies what multicast addresses are available for transmission is housed in yet another part of the TRING system table, at offset `0B80`. Each entry is 8 bytes, and corresponds to the multicast node numbers `09F0`–`09FF`. and an example from node `0614` is as follows:

```
105B80     C000  0000  0000  0000
105B88     C000  0004  0000  0000
```

```
105B90     C000  0000  0000  0000
105B98     C000  0000  0000  0000
105BA0     C000  0000  0000  0000
105BA8     C000  0000  0000  0000
105BB0     C000  0000  0000  0000
105BB8     C000  0000  0000  0000
105BC0     C000  0000  0000  0000
105BC8     C000  0004  0000  0000
105BD0     C000  0400  0000  03D2
105BD8     C000  0000  0000  0000
105BE0     C000  CC49  4E41  0000
105BE8     C000  C45A  524F  0000
105BF0     C000  0000  0000  0000
105BF8     FFFF  FFFF  FFFF  3A2D
```

The first 6 bytes of each entry specifies a token ring group address or a group functional address or a broadcast address. The last word is simply a diagnostic counter, incremented when a frame is sent to the corresponding multicast address.

Note that the entries for and `09F1` and `09F9` both have the standard IP multicast hardware address specified. Node `09F1` is used to target alarm messages via IP multicast. Node `09F9` is used to target name look-up requests or multi-node data requests via IP multicasting. Nodes `09FC` and `09FD` target token ring group addresses `'LINA'` and `'DZRO'` respectively. Node `09FF` targets pure broadcast, as for ARP requests.

One more table relates to the above. It is the third and fourth words in the PAGEM system table, based in token ring nodes at `00102000`. In node `0614`, as an example, we may find the following:

`102004`	`09F9`	Name look-ups and multi-node data requests
`102006`	`09F1`	Alarm messages

The first target node# is used for name look-ups and multi-node requests. The second is used for alarm messages.