Compact GVRP

Abstract

The GARP VLAN Registration Protocol, GVRP, is defined in IEEE Std. 802.1Q-1998. Compact GVRP is a modification to GVRP that reduces the number of packets required to transmit the state of a port. In addition, Compact GVRP makes provision for rapidly pruning VLANs from point-to-point links without use of Leave timers or the Leave All garbage collection message.

1.0 Purpose

1.1 The GVRP Problem

(The reader must become familiar with [1], Clause 12, in order to understand this document.)

In a VLAN bridged network, for a given VLAN ID, there may be only a very few access ports, attached to even few bridges, on which that VLAN is enabled. For this VLAN, it is desirable for a broadcast, multicast, or unknown unicast frame (all of which may be called "floods") from one of those ports to be sent to or through only those bridges that are required to reach the other access ports in that VLAN. Sending floods for that VLAN to all bridges in the network is wasteful of bandwidth. A standard protocol, GVRP[1][2], has been defined that enables bridges to limit the spread of floods based on VLAN ID.

In the average enterprise network, the highest volume of floods comes from multicast streams, which can be limited either by snooping IGMP packets or by the standard GMRP protocol[1]. These protocols are based on limiting the spread of multicast destination MAC addresses. The other types of flooding do not threaten the network's bandwidth, so limiting them is not important. It is rare for and enterprise network to employ more than 100-200 VLANs. As a result, GVRP is very seldom used in the enterprise.

In the Metro Ethernet Provider world, the use of thousands of VLANs is likely. The average Provider VLAN will have only 2-3 exit ports, and the most Providers will neither wish to nor need to filter traffic by destination MAC address. GVRP (or another pruning protocol) will therefore be very important.

1.2 The Problem with GVRP

Unfortunately, it takes 4 octets in a GVRP PDU (Protocol Data Unit) to express the state of each VLAN on a bridge port. In order to transmit all 4094 P-VLAN states on one port, 12 PDUs are required. The processing required when the provider's network undergoes a spanning tree topology change thus becomes very high, as several full transmissions of each port's state will usually be required before GVRP converges. GVRP requires a Registrar state machine and an Applicant state machine for each of the 4094 VLANs. Each Registrar state machine includes a timer (the leavetimer of [1]). Whenever multiple PDUs are required to transmit one station's state, each Applicant state machine also needs its own timer (the jointimer).

The transmission of each VLAN's state as a separate GVRP data registration, the exact parallel with the other current GARP application, GMRP [1], and the lack or need for this protocol (until now) has encouraged vendors to produce a naive but easily-written implementation of the protocol. The end result of this situation is that a bridge CPU can easily become completely bogged down with GVRP processing when the number of VLANs reaches a level around 300-1000 VLANs, and supporting more VLANs becomes impossible.

1.3 An Obvious Solution

One part of the solution to the GVRP problem is the number of state machines required. GVRP requires one Transmit PDU state machine and one LeaveAll state machine, plus any number up to 4094 copies each of the Applicant and Registrar state machines, for a total of 8196 state machines. All of the Applicant state machines could be combined into a single Applicant Engine, and all of the Registrar machine could similarly be combined. Each of these aggregated engines can operate on bit vectors which encode all the VLANs' states.

The timer per VLAN for the 4094 Applicant state machines can be reduced to a single Applicant timer if all of a port's 4094 state messages could be packed into one PDU. Not counting the LeaveAll message, which applies to all registrations for a given application (VLANs), GARP can transmit 4 state messages for a given attribute (VLAN) that the Applicant state machine can transmit: JoinIn, JoinEmpty, Empty, and Leave. The obvious way to shorten GARP frames is simply to replace the list of messages in a GVRP PDU with a vector of 4094 2-bit message numbers. Unfortunately, this neglects the fact that, in standard GVRP, not every VLAN's Applicant state machine needs to send a message. Since in a vector, every Applicant state machine has a message, a fifth message is needed, the "In" message mentioned in [1]. It means, "I have registered this value, but have not joined."

One could encode 4094 VLANs times 5 possible messages per VLAN into 1189 octets by converting a 4094-digit number in base 5 to binary. This approach, however, would require quite a bit of computation in order to extract the 4094 individual messages from the encoded 1189-octet integer. A better approach is to encode 6 messages into each 16 bits. This leads to 1366 octets to encode 4094 messages, which can fit (with the other parts of the GVRP PDU) into one Ethernet frame. Of course, packing 3 messages into 1 octet or 12 messages into 4 octets, etc., is equally possible.

There are still 4094 leave timers in the 4094 Registrar state machines. These will be dealt with in a later version of this document.

1.4 Scope

The aim of Compact GVRP is to make the minimum possible changes to the GVRP specification, while reducing the maximum number of PDUs required to transmit a port's GVRP state from twelve to one.

2.0 Definitions

Applicant Only: An option exists in GVRP for a device to run GVRP in Applicant Only mode. This is appropriate to a device, other that a bridge, which needs to receive certain VLANs, but does not serve as a relay point for VLANs needed by other devices. The most relevant example of such a device is a router. It would use GVRP or Compact GVRP to signal to a bridge the VLANs it services.

Compact Mode: A mode of operation of a Compact GVRP-capable port in which only Standard GVRP PDUs are issued. Compact Mode is further subdivided into Fast Compact Mode and Slow Compact Mode.

Compatible Mode: A mode of operation of a Compact GVRP-capable port in which only Standard GVRP PDUs are issued, but with an additional Negotiation Message indicating that the transmitting device is Compact GVRP-capable. A given port is always either in Compatible Mode or Compact Mode.

Device Identifier: A 6-octet globally unique value identifying the device transmitting a Compact GVRP PDU. This value must be a unicast MAC address belonging to the transmitting device. The Device Identifier is one component of a Source Identifier.

Device Sub-Identifier: A 2-octet value which may be used by the transmitting device to distinguish among multiple ports on which Compact GVRP PDUs are transmitted. The Device Sub-Identifier is one component of a Source Identifier.

Fast Compact Mode: A mode of operation of a Compact GVRP-capable port in which only Negotiation Messages are issued. When a Compact GVRP PDU is received, VLANs may be unregistered immediately, without waiting for a timeout. Fast Compact Mode, with Slow Compact Mode, is one of two subdivisions of Compact Mode.

JustKidding Message: A new GVRP Message used to trick an non-Compact GVRP capable device into responding. Always accompanied by a LeaveAll message.

message_vector: A vector of 683 2-octet integers. Each integer encodes 5 messages from 6 VLANs' Applicant state machines, starting with VLAN 0.

Negotiation Message: A new GVRP Message used to signal the ability of a port to transmit Compact GVRP PDUs. The message contains the sending port's Source Identifier.

Port: For the purposes of this document, the word, "port," is used in place of the word which would be more correct in terms of IEEE 802 standards, "MAC". "Port" is more understandable to most readers.

port_partner: A per-port variable containing a Source Identifier.

Slow Compact Mode: A mode of operation of a Compact GVRP-capable port in which operates the same as standard GVRP, except that a single Compact GVRP PDU can encode the states of all 4094 VLANs' state machines. Slow Compact Mode, with Fast Compact Mode, is one of two subdivisions of Compact Mode.

Source Identifier: The 8 octet value which identifies the source of Compact GVRP PDUs. The Source Identifier consists of a Device Identifier and a Device Sub-Identifier.

Vector Message: A new GVRP message containing the message_vector.

3.0 Protocol Specification

3.1 Basic differences between GVRP and Compact GVRP

Compact GVRP uses the same state machines as standard GVRP, except in Fast Compact Mode. The main differences between standard and Compact GVRP are in the PDUs transmitted. Compact GVRP accomplishes the reduction PDU count by reducing the number of bits transmitted in a PDU to:

1. A Vector Message containing a message_vector of 683 2-octet integers, carrying one of the 5 GVRP messages JoinEmpty, JoinIn, LeaveEmpty, Empty, or In, from each of the 4094 possible VLAN-IDs.
2. A Negotiation Message, containing an 8-octet Source Identifier.

The Negotiation Message makes it possible for Compact GVRP-capable ports to discover each other so that they can operate in Compact Mode. The Negotiation Message is also added to standard GVRP PDUs transmitted by a Compact GVRP-capable port. This allows Compact GVRP-capable devices to discover each other, to discover whether there are two of them or more, and allows them to interoperate with devices that run only Standard GVRP.

It must be noted that these two new messages cannot be accommodated by the current GARP PDU format, and thus cannot be transmitted by a GVRP PDU, which is a particular case of a GARP PDU. However, the above information can be mapped to the information that is transmitted in the current GVRP PDU, so can be made compatible with the GARP state machines as used in standard GVRP.

3.2 Compatibility Negotiation

A Compact GVRP-capable port is either in "Compatible Mode" or "Compact Mode". If in Compact Mode, the port may further be in either "Fast Compact Mode" or "Slow Compact Mode." A Compact GVRP port is initialized to Compatible Mode. In this mode, it transmits only standard GVRP PDU, with a single exception: a new Negotiation Message is included in each PDU. This new Attribute Type indicates that the transmitting device is capable of operating in Compact Mode. The existing IEEE Std. 802.1D-2003 Clause 12.11.3.3 specifies that any GARP application must ignore any Attribute Type that is does not understand. Thus, this new Negotiation Message does not confuse an old GVRP implementation. However, the presence of the Negotiation Message allows Compact GVRP-capable devices to discover each other.

If any port on the LAN is not Compact GVRP-capable, and so transmits a standard GVRP PDU that does not contain the Negotiation Message, all Compact GVRP-capable ports revert to Compatible Mode, so that all can communicate. The object of compatibility negotiation is to discover whether there are any devices that are not Compact GVRP-capable, and if there are none, whether there are more than two Compact GVRP devices. If there are only two devices, they can operate in Fast Compact Mode, which is faster to converge than the other modes.

In all cases, if switching operational modes causes a change in the state machines in use, the switch takes place before any of the VLAN messages in the GVRP are processed. The state in the new state machine is the one corresponding to the state in the old machine according to Table NWF. The rules for transitioning between “Compatible Mode,” “Fast Compact Mode,” and “Slow Compact Mode” are:

1. All ports are initialized to Compatible Mode. They are reinitialized each time the port transitions from a non-operational to an operational state. (This condition is normally determined by the port’s ifOperState MIB variable.) As part of the initialization process, the port’s port_partner variable is initialized to a value of all 1s.
2. Using the same kind of mechanism as the LeaveAll state machine, only using a timer justkiddingtimer which is much longer than the leavealltimer, the port sends a PDU containing both a JustKidding Message and a LeaveAll Message. This PDU is designed to tease out any non-Compact GVRP-capable ports. It looks like a LeaveAll message to such a port, but the JustKidding Message causes a Compact GVRP-capable port to ignore the PDU. A JustKidding PDU is also sent immediately before the first LeaveAll message sent by any port after initialization. The JustKidding state machine uses a justkiddingtimer that is, by default, 10 times the length of the leavealltimer.
3. Whenever a JustKidding Message is transmitted, a leavetimer in the JustKidding state machine is started. If that leavetimer expires without the receipt of a Standard GVRP PDU which does not contain the Negotiation Message, then the port enters Slow Compact Mode.
4. Upon receipt of a Standard GVRP PDU which does not contain the Negotiation Message, a port enters Compatible Mode, and the port’s port_partner variable is reset to a value of all 1s.
5. All Compact GVRP PDUs are required to carry a Negotiation Message. In a Negotiation Message, a Device Sub-Identifier of all 1s is illegal.
6. Upon receipt of a Negotiation Message, if the port_partner contains all 1s, the PDU’s Source Identifier is stored in the port_partner variable. If not, then the PDU’s Source Identifier is compared to the port’s port_partner variable. If equal, or if the port is in Compatible Mode, the port’s mode of operation is unchanged; else, if unequal and the port is in Fast Compact Mode, the port is set to Slow Compact Mode; else (the port is therefore in Slow Compact Mode), the operational mode does not change.

NOTE: If a non-Compact GVRP-capable device implements only the Registrar state machine, and not an Applicant state machine, or if it implements only one of the Applicant state machines and no Registrar state machine and is not currently a member of a VLAN, then it may not respond to a LeaveAll/JustKidding Message, and thus reveal its presence to the Compact GVRP-capable device(s). If this situation is likely to occur, a Compact GVRP-capable device must be restricted by configuration to operate in Compatible Mode only.

3.3 Compact GVRP PDU Format

A Compact GVRP PDU uses the same Protocol ID as standard GARP. However, two new Attribute Types are use for the two new Compact GVRP messages, the Negotiation Message and the Vector Message. A Compact GVRP PDU can encode all of the VLANs’ Attribute structures into a single GARP message. Each Attribute structure encodes a portion of the (up to) 683 2-octet integers constituting a message_vector. The message formats are shown in Figure 1.

Within each Attribute structure in the Vector Message are an integral number of half-words. Each half-word is in network order (most-significant octet first), and encodes 6 normal GVRP Attribute Events E for VLANs x through $x + 4$ using the formula:

$$\text{half-word} = ((((E_x \cdot 5 + E_{x+1}) \cdot 5 + E_{x+2}) \cdot 5 + E_{x+3}) \cdot 5 + E_{x+4}) \cdot 5 + E_{x+5} \quad (1)$$

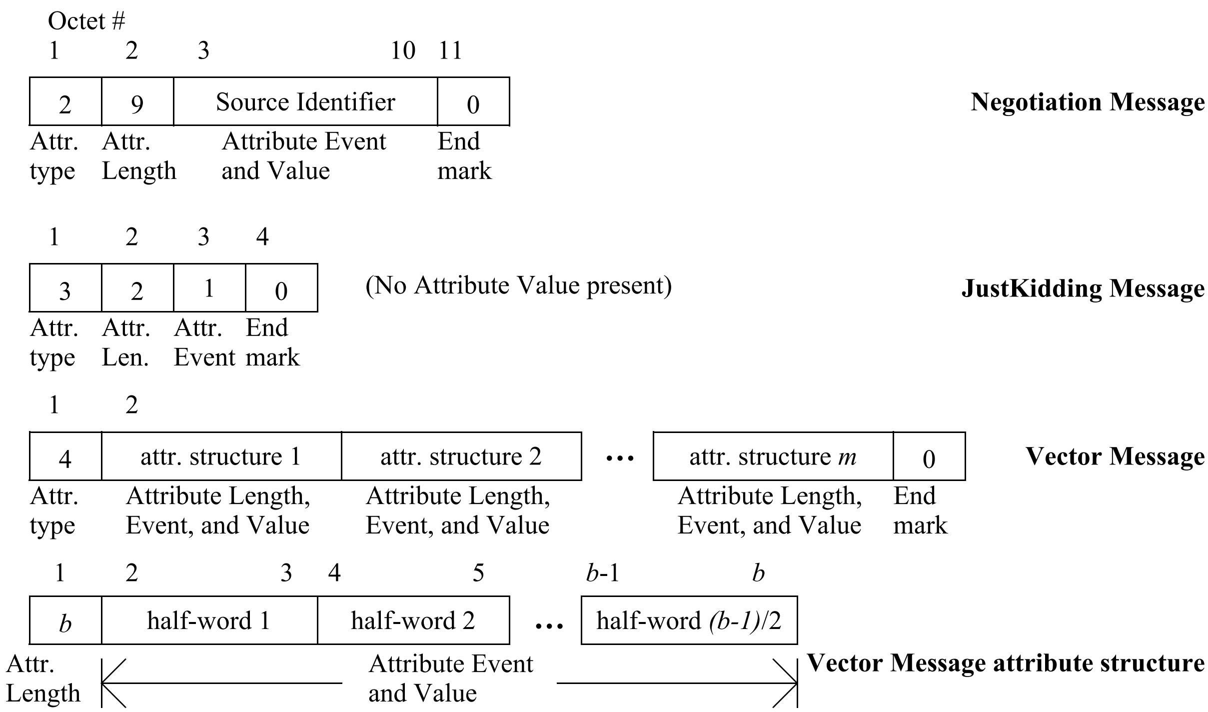
To decode the half-word, one may run a series of divisions:

$$E_{x+5} = rem(\text{half-word}, 5)$$

$$\text{temp} = \text{half-word} / 5$$

$$E_{x+4} = rem(\text{temp}, 5)$$

Figure 1: Compact GVRP Message Formats



$$\text{temp} = \text{temp}/5$$

$$\ldots$$

$$E_{x+1} = rem(\text{temp}, 5)$$

$$E_x = \text{temp}/5$$

where all divisions are integer divisions, and *rem*() is the integer remainder function.

Table 1 shows the encoding of the Attribute Events *E*. No LeaveIn Attribute is encoded, because none need be sent in a GVRP message. The new In Attribute Event is included because, unlike standard GVRP, VLANs' Applicant state machines must often send a message just to fill out the message_vector, when they would not otherwise send one.

Table 1: Attribute Event codes

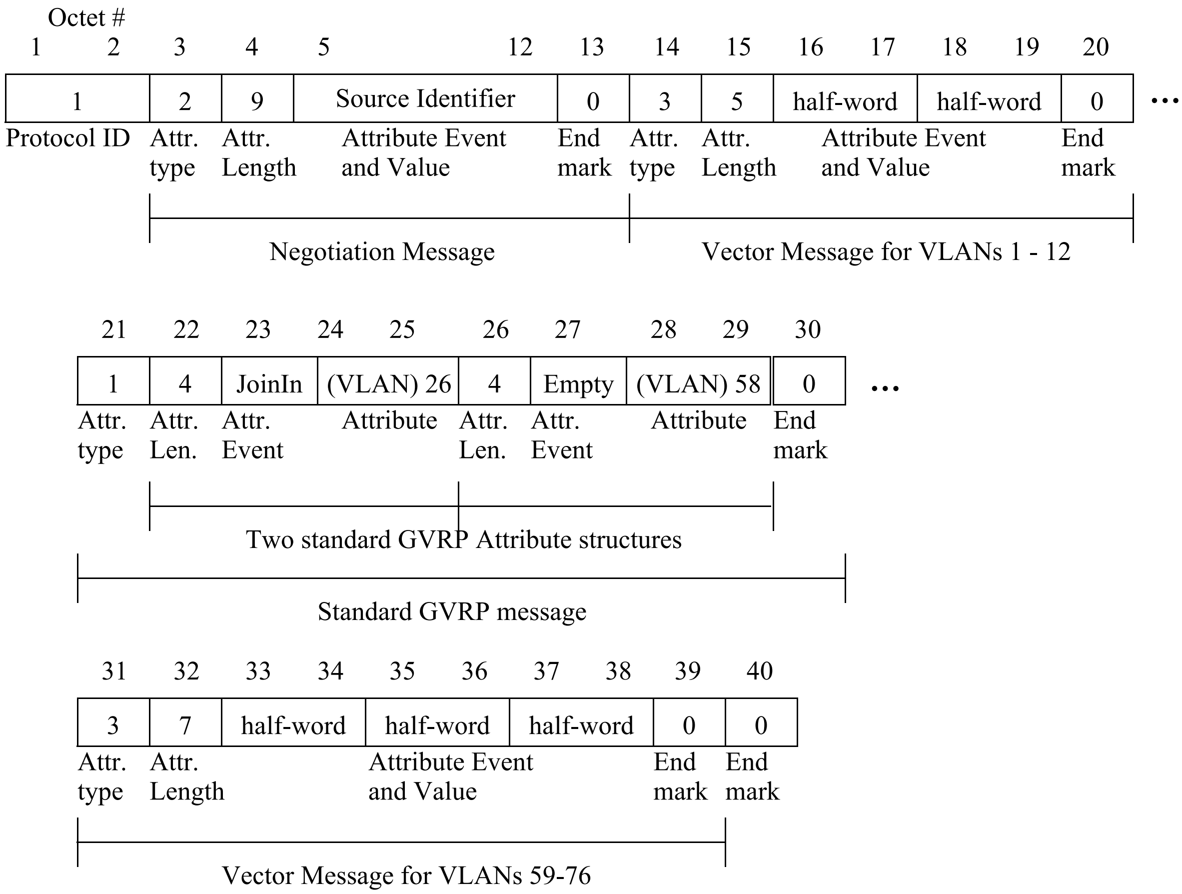
message_vector code value *E*	Attribute Event
0	In
1	JoinEmpty
2	JoinIn
3	LeaveEmpty
4	Empty

NOTE: These Attribute Event codes are not the same codes used in standard GVRP.

The first half-word in the message encodes the Attribute Events for VLANs x+1 through x+6, the second for

VLANs *x*+7 through *x*+12, etc. where *x* is the VLAN ID of the immediately previous VLAN included in the GVRP PDU, or 0, if this is the first message with Attribute Type 1 or 3 in the PDU. The half-words may packed into any number of Attribute structures an a Vector Message, and there may be any number of Vector Messages in a Compact GVRP PDU. At least six Attribute structures are needed to hold all 4094 VLANs' Attribute Events. Standard GVRP messages and Vector Messages may be interspersed. Each Vector Message includes an integral number of half-words, hence a multiple of 6 VLANs' Attribute Events. The last half-word in a Vector Message, therefore, may include Attributes for as many as 5 illegal VLAN IDs. The Attribute Events for VLANs greater than 4094 must be encoded as the "In" event. See Figure 2 for an example of a mixed-format PDU.

Figure 2: Mixed-format GVRP Message



The formats of the Negotiation Message and Vector Message violate the formats specified in [1], because neither has an Attribute Event field in the sense intended by the standard. However, because the Attribute Type at the beginning of the message is unknown to a standard GVRP implementation, and because the position and meaning of the Length fields in the Negotiation Message and Vector Message are correct, a standard GVRP implementation should have no problem in skipping over these unknown messages.

If present, the Negotiation Message must be the first message in a Compact GVRP PDU. Any GVRP PDU which includes one or more Vector Messages but no Negotiation Message is invalid and should be discarded. Messages up to but not including the first Vector Message may or may not be processed by the receiver, in that case.

A Compact GVRP implementation must always transmit every VLAN's state in every Compact GVRP PDU. Upon receipt, the presence of a Negotiation Message in a GVRP PDU implies an "In" message for every VLAN not explicitly mentioned in that PDU.

3.4 Compatible Mode and Slow Compact Mode State Machines

In both Compatible Mode and Slow Compact Mode, the standard GARP state machines can be used unaltered. The "In" message is not in the state machines, even though In messages may be received, because it does not cause a state transition.

3.5 Fast Compact Mode Combined Applicant / Registrar State Machine

The Fast Compact Mode Combined state machine in Table 3 has several new actions and one new event:

sI Action. Send an "In" message.

xJI, xJE, xI, xE Actions. Send a JoinIn, JoinEmpty, In, or Empty message, but only if some other VLAN's state machine has an sX event. If all VLANs are in a state with an xX event, no Compact GVRP PDU is emitted.

rIn Event. "In" message received.

The state machine, itself, is shown in The differences between Table 2 and Table 3 occur because there are only two devices on the LAN. To summarize these differences:

Table 2: Compatible Mode / Slow Compact Mode Applicant State Machine (from 802.1D[1])

		STATE (initial state: VO)										
		VA	**AA**	**QA**	**LA**	**VP**	**AP**	**QP**	**VO**	**AO**	**QO**	**LO**
EVENT	**transmitPDU!**	sJ[E,I] AA	sJ[E,I] QA	-x-	sLE VO	sJ[E,I] AA	sJ[E,I] QA	-x-	-x-	-x-	-x-	sE VO
	rJoinIn	AA	QA	QA	LA	AP	QP	QP	AO	QO	LO	AO
	rJoinEmpty	VA	VA	VA	VO	VP	VP	VP	VO	VO	VO	VO
	rEmpty	VA	VA	VA	LA	VP	VP	VP	VO	VO	VO	VO
	rLeaveIn	VA	VA	VA	LA	VP	VP	VP	LO	LO	LO	LO
	rLeaveEmpty	VP	VP	VP	VO	VP	VP	VP	LO	LO	LO	VO
	LeaveAll	VP	VP	VP	VO	VP	VP	VP	LO	LO	LO	VO
	ReqJoin	-x-	-x-	-x-	VA	-x-	-x-	-x-	VP	AP	QP	VP
	ReqLeave	LA	LA	LA	-x-	VO	AO	QO	-x-	-x-	-x-	-x-

1. Except for the transmitPDU! event, there is no difference among Very anxious, Anxious, and Quiet Applicants.
2. There is no need for Passive Applicants, because there is no need to avoid sending too many Leave messages. Therefore, all Applicants are Active.
3. An Applicant sends up to two Compact GVRP PDUs to make sure that the other device's Registrar's state agrees with the first device's Applicant's state. If an Applicant finds that its state differs from the other Registrar, it sends two PDUs to correct it. If a Registrar finds that its state differs from the other device's Applicant, it sends one PDU to acknowledge the change. In terms of the diagrams:
4. If a message is received showing a Registrar state that conflicts with the receiver's Applicant state, or if a ReqJoin or ReqLeave occur that change my Applicant between Member (Active) or Observer, then the Applicant becomes Very anxious; **else**
   if a message is received showing an Applicant state that conflicts with the receiver's Registrar state, then the receiver's Applicant becomes Anxious; **else**

Table 3: Fast Compact Mode Combined Attribute / Registrar State Machine

		STATE											
	Applicant:	**VA**	**AA**	**QA**	**VO**	**AO**	**QO**	**VA**	**AA**	**QA**	**VO**	**AO**	**QO**
	Registrar:	**IN**						**MT**					
EVENT	**transmitPDU!**	sJI AA IN	sJI QA IN	xJI QA IN	sI AO IN	sI QO IN	xI QO IN	sJE AA MT	sJE QA MT	xJI QA MT	sE AO MT	sE QO MT	xE QO MT
	rJoinIn	QA IN	QA IN	QA IN	VO IN	VO IN	VO IN	AA IN	AA IN	AA IN	VO IN	VO IN	VO IN
	rJoinEmpty	VA IN	VA IN	VA IN	QO IN	QO IN	QO IN	VA IN	VA IN	VA IN	AO IN	AO IN	AO IN
	rEmpty	VA MT	VA MT	VA MT	AO MT	AO MT	AO MT	VA MT	VA MT	VA MT	QO MT	QO MT	QO MT
	LeaveIn	AA MT	AA MT	AA MT	VA MT	VA MT	VA MT	QA MT	QA MT	QA MT	VO MT	VO MT	VO MT
	rLeaveEmpty	VA MT	VA MT	VA MT	AA MT	AA MT	AA MT	VA MT	VA MT	VA MT	QO MT	QO MT	QO MT
	rIn	AA MT	AA MT	AA MT	VA MT	VA MT	VA MT	QA MT	QA MT	QA MT	VO MT	VO MT	VO MT
	LeaveAll	VA MT	VA MT	VA MT	AA MT	AA MT	AA MT	VA MT	VA MT	VA MT	QO MT	QO MT	QO MT
	ReqJoin	-x-	-x-	-x-	VA IN	VA IN	VA IN	-x-	-x-	-x-	VA MT	VA MT	VA MT
	ReqLeave	VO IN	VO IN	VO IN	-x-	-x-	-x-	VO MT	VO MT	VO MT	-x-	-x-	-x-

the receiver's Applicant becomes Quiet.

5. No distinction need be made between LeaveEmpty and Empty, nor between LeaveIn and In.
6. In effect, with two devices, they merely trade their Member vs. Observer and Registered vs. Not Registered states until they agree.

NWF: Do we need xI for slow mode machine? Sending all In messages needs no Vector Message.

References

[1] IEEE Std. 802.1D-2003; "Media Access Control (MAC) Bridges"; http://p8021:go_wildcats@www.ieee802.org/1/files/private/d-rev-drafts/d3/802-1D-D3.pdf

[2] IEEE Std. 802.1Q-1998; "Virtual Bridged Local Area Networks".