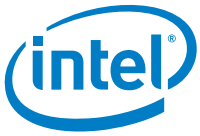


6.4.3.10 LAN 0/1 Ethernet MAC Address 1 MMAL/H1 — Offset 0x0C:0x0E

Same structure as LAN0 Ethernet MAC Address 0. Loaded to MMAL[1], MMAH[1].

6.4.3.11 LAN 0/1 Ethernet MAC Address 2 MMAL/H2 — Offset 0x0F:0x11

Same structure as LAN0 Ethernet MAC Address 0. Loaded to MMAL[2], MMAH[2].

6.4.3.12 LAN 0/1 Ethernet MAC Address 3 MMAL/H3 — Offset 0x12:0x14

Same structure as LAN0 Ethernet MAC Address 0. Loaded to MMAL[3], MMAH[3].

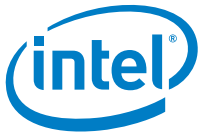
6.4.3.13 LAN 0/1 UDP Flexible Filter Ports 0:15 (MFUTP Registers) - Offset 0x15:0x24

Offset	Bits	Description	Reserved
0x15	15:0	LAN UDP Flexible Filter Value Port0.	
0x16	15:0	LAN UDP Flexible Filter Value Port1.	
0x17	15:0	LAN UDP Flexible Filter Value Port2.	
0x18	15:0	LAN UDP Flexible Filter Value Port3.	
0x19	15:0	LAN UDP Flexible Filter Value Port4.	
0x1A	15:0	LAN UDP Flexible Filter Value Port5.	
0x1B	15:0	LAN UDP Flexible Filter Value Port6.	
0x1C	15:0	LAN UDP Flexible Filter Value Port7.	
0x1D	15:0	LAN UDP Flexible Filter Value Port8.	
0x1E	15:0	LAN UDP Flexible Filter Value Port9.	
0x1F	15:0	LAN UDP Flexible Filter Value Port10.	
0x20	15:0	LAN UDP Flexible Filter Value Port11.	
0x21	15:0	LAN UDP Flexible Filter Value Port12.	
0x22	15:0	LAN UDP Flexible Filter Value Port13.	

Offset	Bits	Description	Reserved
0x23	15:0	LAN UDP Flexible Filter Value Port14.	
0x24	15:0	LAN UDP Flexible Filter Value Port15.	

6.4.3.14 LAN 0/1 VLAN Filter 0 — 7 (MAVTV Registers) - Offset 0x25:0x2C

Offset	Bits	Description	Reserved
0x25	15:12	Reserved	
0x25	11:0	LAN 0/1 VLAN filter 0 value.	
0x26	15:12	Reserved	
0x26	11:0	LAN 0/1 VLAN filter 1 value.	
0x27	15:12	Reserved	
0x27	11:0	LAN 0/1 VLAN filter 2 value.	
0x28	15:12	Reserved	
0x28	11:0	LAN 0/1 VLAN filter 3 value.	
0x29	15:12	Reserved	
0x29	11:0	LAN 0/1 VLAN filter 4 value.	
0x2A	15:12	Reserved	
0x2A	11:0	LAN 0/1 VLAN filter 5 value.	
0x2B	15:12	Reserved	
0x2B	11:0	LAN 0/1 VLAN filter 6 value.	
0x2C	15:12	Reserved	
0x2C	11:0	LAN 0/1 VLAN filter 7 value.	



6.4.3.15 LAN 0/1 Manageability Filters Valid (MFVAL LSB) — Offset 0x2D

Bits	Name	Default	Description	Reserved
15:8	VLAN		Indicates if the VLAN filter registers (MAVTV) contain valid VLAN tags. Bit 8 corresponds to filter 0, etc.	
7:4	Reserved			
3:0	MAC		Indicates if the MAC unicast filter registers (MMAH, MMAL) contain valid Ethernet MAC Addresses. Bit 0 corresponds to filter 0, etc.	

6.4.3.16 LAN 0/1 Manageability Filters Valid (MFVAL MSB) — Offset 0x2E

Bits	Name	Default	Description	Reserved
15:12	Reserved			
11:8	IPv6		Indicates if the IPv6 address filter registers (MIPAF) contain valid IPv6 addresses. Bit 8 corresponds to address 0, etc. Bit 11 (filter 3) applies only when IPv4 address filters are not enabled (MANC.EN_IPv4_FILTER=0).	
7:4	Reserved		Reserved.	
3:0	IPv4		Indicates if the IPv4 address filters (MIPAF) contain a valid IPv4 address. These bits apply only when IPv4 address filters are enabled (MANC.EN_IPv4_FILTER=1).	

6.4.3.17 LAN 0/1 MANC value LSB (LMANC LSB) — Offset 0x2F

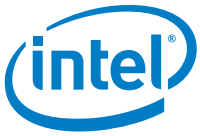
Bits	Name	Default	Description	Reserved
15:0	Reserved	0x0	Reserved.	

6.4.3.18 LAN 0/1 MANC Value MSB (LMANC MSB) — Offset 0x30

Bits	Name	Default	Description	Reserved
15:9	Reserved		Reserved.	
8	Enable IPv4 Address Filters		This bit is loaded to the EN_IPv4_FILTER bit in the MANC register.	
7	Enable Xsum Filtering to MNG		This bit is loaded to the EN_XSUM_FILTER bit in the MANC register.	
6	VLAN MNG Filtering		This bit is loaded to the *Bypass VLAN* bit in the MANC register.	
5	Enable MNG Packets to Host Memory		This bit is loaded to the EN_MNG2HOST bit in the MANC register.	
4:0	Reserved		Reserved.	

6.4.3.19 LAN 0/1 Receive Enable 1 (LRXEN1) — Offset 0x31

Bits	Name	Default	Description	Reserved
15:8	Receive Enable Byte 12		BMC SMBus slave address.	
7	Enable BMC Dedicated MAC			
6	Reserved		Reserved. Must be set to 1b.	
5:4	Notification Method		00b = SMBus alert. 01b = Asynchronous notify. 10b = Direct receive. 11b = Reserved.	
3	Enable ARP Response			
2	Enable Status Reporting			
1	Enable Receive All			
0	Enable Receive TCO			



6.4.3.20 LAN 0/1 Receive Enable 2 (LRXEN2) — Offset 0x32

Bits	Name	Default	Description	Reserved
15:8	Receive Enable byte 14	0x0	Alert value.	
7:0	Receive Enable byte 13	0x0	Interface data.	

6.4.3.21 LAN 0/1 MANC2H Value (LMANC2H LSB) — Offset 0x33

Bits	Name	Default	Description	Reserved
15:8	Reserved			
7:0	Host Enable		When set, indicates that packets routed by the manageability filters to the manageability block are also sent to the host. Bit 0 corresponds to decision rule 0, etc.	

6.4.3.22 LAN 0/1 MANC2H Value — LMANC2H MSB - Offset 0x34

Bits	Name	Default	Description	Reserved
15:0	Reserved	0x0	Reserved	

6.4.3.23 Manageability Decision Filters — MDEF0 (1) - Offset 0x35

Bits	Name	Default	Description	Reserved
15:0	MDEF0_L		Loaded to 16 LS bits of MDEF[0] register.	

6.4.3.24 Manageability Decision Filters — MDEF0 (2) - Offset 0x36

Bits	Name	Default	Description	Reserved
15:0	MDEF0_M		Loaded to 16 MS bits of MDEF[0] register.	

6.4.3.25 Manageability Decision Filters — MDEF0 (3) - Offset 0x37

Bits	Name	Default	Description	Reserved
15:0	MDEFEXT0_L		Loaded to 16 LS bits of MDEF_EXT[0] register.	

6.4.3.26 Manageability Decision Filters — MDEF0 (4) - Offset 0x38

Bits	Name	Default	Description	Reserved
15:0	MDEF0EXT_M		Loaded to 16 MS bits of MDEF_EXT[0] register.	

6.4.3.27 Manageability Decision Filters — MDEF1-6 (1-4) - Offset 0x39:0x50

Same as words 0x035...0x38 for MDEF[1] and MDEF_EXT[1]...MDEF[6] and MDEF_EXT[6]

6.4.3.28 Manageability Ethertype Filter 0.1 — METF0 (1) - Offset 0x51

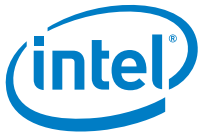
Bits	Name	Default	Description	Reserved
15:0	METF0_L		Loaded to 16 LS bits of METF[0] register.	

6.4.3.29 Manageability Ethertype Filter 0.2 — METF0 (2) - Offset 0x52

Bits	Name	Default	Description	Reserved
15:0	METF0_M		Loaded to 16 MS bits of METF[0] register (reserved bits in the METF registers should be set in the EEPROM to the register's default values).	

6.4.3.30 Manageability Ethertype Filter 1...3 (1 and 2) — METF1:3 - Offset 0x53:0x58

Same as words 0x51 and 0x52 for METF[1]...METF[3] registers.



6.4.3.31 ARP Response IPv4 Address 0 (LSB) — Offset 0x59

Bits	Name	Default	Description	Reserved
15:0			ARP Response IPv4 Address 0, Byte 1 (firmware use).	
7:0			ARP Response IPv4 Address 0, Byte 0 (firmware use).	

6.4.3.32 ARP Response IPv4 Address 0 (MSB) — Offset 0x5A

Bits	Name	Default	Description	Reserved
15:8			ARP Response IPv4 Address 0, Byte 3 (firmware use).	
7:0			ARP Response IPv4 Address 0, Byte 2 (firmware use).	

6.4.3.33 LAN 0/1 IPv6 Address 0 (n=0...7) (MIPAF.IPV6ADDR0) — Offset 0x5B:0x62

Bits	Name	Default	Description	Reserved
15:0			Loaded to MIPAF registers IPV6ADDR0: Dword offset 'n'/2 to the lower 16 bits for even 'n' and upper 16 bits for odd 'n'. For 'n' = 0...7.	

6.4.3.34 LAN 0/1 IPv6 Address 1 (MIPAF.IPV6ADDR1) — Offset 0x63:0x6A

Same structure as LAN 0/1 IPv6 Address 0.

6.4.3.35 LAN 0/1 IPv6 Address 2 (MIPAF) — Offset 0x6B-0x72

Same structure as LAN 0/1 IPv6 Address 0.

6.4.4 Sideband Configuration Module

This module is pointed to by global offset 0x06 of the manageability control table.

6.4.4.1 Section Header — Offset 0x0

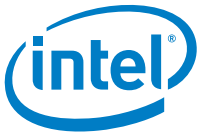
Bits	Name	Default	Description	Reserved
15:8	Block CRC8	0x0		
7:0	Block Length	0x0	Section length in words.	

6.4.4.2 SMBus Maximum Fragment Size — Offset 0x01

Bits	Name	Default	Description	Reserved
15:0	Max Fragment Size	0x0	SMBus Maximum Fragment Size (bytes)	

6.4.4.3 SMBus Notification Timeout and Flags — Offset 0x02

Bits	Name	Default	Description	Reserved
15:8	SMBus Notification Timeout (ms)			
7:6	SMBus Connection Speed		00b = Standard SMBus connection. 01b = Reserved. 10b = Reserved. 11b = Reserved.	
5	SMBus Block Read Command		0b = Block read command is 0xC0. 1b = Block read command is 0xD0.	
4	SMBus Addressing Mode		0b = Single address mode. 1b = Dual address mode.	
3	Reserved		Reserved.	
2	Disable SMBus ARP Functionality			
1	SMBus ARP PEC			
0	Reserved		Reserved.	



6.4.4.4 SMBus Slave Addresses — Offset 0x03

Bits	Name	Default	Description	Reserved
15:9	SMBus 1 Slave Address		Dual address mode only.	
8	Reserved		Reserved.	
7:1	SMBus 0 Slave Address			
0	Reserved		Reserved.	

6.4.4.5 Fail-Over Register (Low Word) — Offset 0x04

Bits	Name	Default	Description	Reserved
15:12	Gratuitous ARP Counter			
11:10	Reserved		Reserved.	
9	Enable Teaming Fail-Over on DX			
8	Remove Promiscuous on DX			
7	Enable MAC Filtering			
6	Enable Repeated Gratuitous ARP			
5	Reserved		Reserved.	
4	Enable Preferred Primary			
3	Preferred Primary Port			
2	Transmit Port			
1:0	Reserved		Reserved.	

6.4.4.6 Fail-Over Register (High Word) — Offset 0x05

Bits	Name	Default	Description	Reserved
15:8	Gratuitous ARP Transmission Interval (seconds)			
7:0	Link Down Fail-Over Time			

6.4.4.7 NC-SI Configuration Offset 0x06

Bits	Name	Default	Description	Reserved
15:14	Reserved		Reserved.	
13	Enable Channel Swap	0b	0b = Legacy mode — NCSI channel is aligned with physical port if LFS bit enabled. 1b = NCSI channels swapped with the physical port if LFS bit enabled.	
12:11	Reserved		Reserved.	
10	Reserved	0b	Must be 0.	
9	NC-SI HW Arbitration Enable	0b	0b = Not supported.Must be set to 0b. 1b = Supported.	
8	NC-SI HW-based Packet Copy Enable	1b	0b = Disable. 1b = Enable.	
7:5	Package ID	000b		
4:0	Reserved	0x0	Must be 0.	

6.4.4.8 Reserved Words — Offset 0x07 - 0x0C

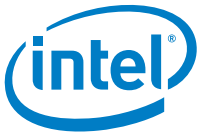
Reserved for future use.

6.4.5 Flexible TCO Filter Configuration Module

This module is pointed to by global offset 0x07 of the manageability control section.

6.4.5.1 Section Header — Offset 0x0

Bits	Name	Default	Description	Reserved
15:8	Block CRC8	0x0		
7:0	Block Length	0x0	Section length in words.	



6.4.5.2 Flexible Filter Length and Control — Offset 0x01

Bits	Name	Default	Description	Reserved
15:8	Flexible Filter Length (bytes)			
7:5	Reserved		Reserved.	
4	Last Filter			
3:2	Filter Index (0-3)			
1	Apply Filter to LAN 1			
0	Apply Filter to LAN 0			

6.4.5.3 Flexible Filter Enable Mask — Offset 0x02 - 0x09

Bits	Name	Default	Description	Reserved
15:0	Flexible Filter Enable Mask			

6.4.5.4 Flexible Filter Data — Offset 0x0A - Block Length

Bits	Name	Default	Description	Reserved
15:0	Flexible Filter Data			

Note: This section loads all of the flexible filters, The control + mask + filter data are repeatable as the number of filters. Section length in offset 0 is for all filters.

6.4.6 NC-SI Microcode Download Module

This module is pointed to by global offset 0x09 of the manageability control table.

6.4.6.1 Patch Data Size — Offset 0x0

6.4.6.2 Rx and Tx Code Size — Offset 0x1

Bits	Name	Default	Description	Reserved
15:8	Rx Code Length	0x0	Rx Code length in Dwords.	
7:0	Tx Code Length	0x0	Tx Code length in Dwords.	

6.4.6.3 Download Data — Offset 0x2 - Data Size

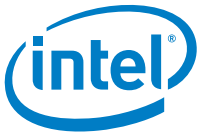
Bits	Name	Default	Description	Reserved
15:8	Download Data	0x0	Download data.	

6.4.7 NC-SI Configuration Module

This module is pointed to by global offset 0x0A of the manageability control table.

6.4.7.1 Section Header — Offset 0x0

Bits	Name	Default	Description	Reserved
15:8	Block CRC8	0x0		
7:0	Block Length	0x0	Section length in words.	



6.4.7.2 Rx Mode Control1 (RR_CTRL[15:0]) — Offset 0x1

Bits	Name	Default	Description	Reserved
15:8	Reserved		Set to 0x0.	
7:4	Reserved		Reserved.	
3	NC-SI Speed		When set, the NC-SI MAC speed is 100 Mb/s. When reset, NC-SI MAC speed is 10 Mb/s.	
2	Receive Without Leading Zeros		If set, packets without leading zeros (J/K/ symbols) between TXEN assertion and TXD the first preamble byte can be received.	
1	Clear Rx Error		Should be set when the Rx path is stuck because of an overflow condition.	
0	NC-SI Loopback Enable		When set, enables NC-SI Tx-to-Rx loop. All data that is transmitted from NC-SI is returned to it. No data is actually transmitted from NC-SI.	

6.4.7.3 Rx Mode Control2 (RR_CTRL[31:16]) — Offset 0x2

Bits	Name	Default	Description	Reserved
15:0	Reserved	0x0		

6.4.7.4 Tx Mode Control1 (RT_CTRL[15:0]) — Offset 0x3

Bits	Name	Default	Description	Reserved
15:3	Reserved		Set to 0x0.	
2	Transmit With Leading Zeros		When set, sends leading zeros (J/K/ symbols) from CRS_DV assertion to the start of preamble (PHY mode). When de-asserted, does not send leading zeros (MAC mode).	
1	Clear Tx Error		Should be set when Tx path is stuck because of an underflow condition. Cleared by hardware when released.	
0	Enable Tx Pads		When set, the NC-SI TX pads are driving. Otherwise, they are isolated.	

6.4.7.5 Tx Mode Control2 (RT_CTRL[31:16]) — Offset 0x4

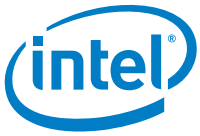
Bits	Name	Default	Description	Reserved
15:0	Reserved	0x0	Set to 0x0.	

6.4.7.6 MAC Tx Control Reg1 (TxCntrlReg1 (15:0]) — Offset 0x5

Bits	Name	Default	Description	Reserved
15:7	Reserved	0x0	Set to 0x0.	
6	NC-SI_enable		Enable the MAC internal NC-SI mode of operation (disables external NC-SI gasket).	
5	Two_part_deferral		When set, performs the optional two part deferral.	
4	Append_fcs		When set, computes and appends the FCS on Tx frames.	
3	Pad_enable		Pad the TX frames, which are less than the minimum frame size.	
2:1	Reserved		Reserved.	
0	Tx_ch_en		Tx Channel Enable. This bit can be used to enable the Tx path of the MAC. This bit is for debug only and the recommended way to enable the Tx path is via the RT_UCTL_CTRL.TX_enable bit.	

6.4.7.7 MAC Tx Control Reg2 (TxCntrlReg1 (31:16]) — Offset 0x6

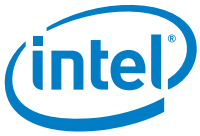
Bits	Name	Default	Description	Reserved
15:0	Reserved		Reserved. Should be set to 0x0.	



6.4.7.8 NC-SI Settings — Offset 0x7

Bits	Name	Default	Description	Reserved
15:9	Reserved	0x0	Set to 0x0.	
8:7	RMM Out Slew Rate	01b	Configuration of the NC-SI out slew-rate control. 00b = Slowest 01b = Slow 10b = Fast 11b = Fastest	
6:1	RMM Out Buffer Strength	011111b	Configuration of the NC-SI out buffer strength. 000001b = 2 mA 000011b = 4 mA 000111b = 6 mA 001111b = 8 mA 011111b = 10 mA 111111b = 12 mA	
0	Reserved	0b	Set to 0b.	

NOTE: ***This page intentionally left blank.***



7.0 Inline Functions

7.1 Receive Functionality

Packet reception consists of:

- Recognizing the presence of a packet on the wire
- Performing address filtering
- DMA queue assignment
- Storing the packet in the receive data FIFO
- Transferring the data to assigned receive queues in host memory
- Updating the state of a receive descriptor.

A received packet goes through three stages of filtering as depicted in Figure 7-1. The Figure describes a switch-like structure that is used in virtualization mode to route packets between the network port (top of drawing) and one of many virtual ports (bottom of drawing), where each virtual port might be associated with a Virtual Machine (VM), an IOVM, a VMM, or the like. The three stages are:

1. First stage — Ensure that the packet should be received by the port. This is done by a set of L2 filters and is described in detail in Section 7.1.1.
2. Second stage — This stage is specific to virtualization environments and defines the virtual ports (called pools in this document) that are the targets for the Rx packet. A packet can be associated with any number of ports/pools and the selection process is described in Section 7.1.2.2.
3. Third stage — A receive packet that successfully passed the Rx filters is associated with one of many receive descriptor queues as described in this section.

In addition to the filtering rules, a packet must also meet the following criteria:

1. Normally, only good packets are received (packets with none of the following errors: Under Size Error, Over Size Error, Packet Error, Length Error and CRC Error). However, if the store-bad-packet bit is set (FCTRL.SBP), then bad packets that don't pass the filter function are stored in host memory. Packet errors are indicated by error bits in the receive descriptor (RDESC.ERRORS). It is possible to receive all packets, regardless of whether they are bad, by setting the promiscuous enables bit and the store-bad-packet bit.
2. Min Packet Size (Runt packets) — Rx packets, smaller than 21 bytes, cannot be posted to host memory regardless of save bad frame setting.

3. Max Packet Size — Any Rx packet posted from the MAC unit to the DMA unit cannot exceed
   15.5 KB.

Note: CRC errors before the SFD are ignored. All packets must have a valid SFD in order to be recognized by the device (even bad packets).

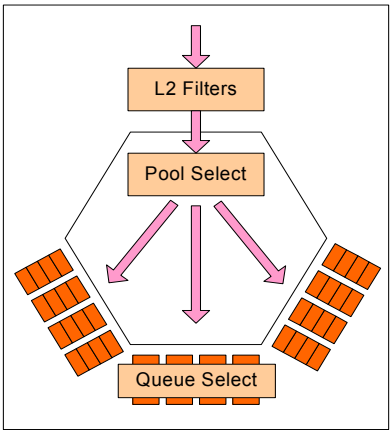


Figure 7-1 Stages in Packet Filtering

7.1.1 Packet Filtering

The receive packet filtering role is to determine which of the incoming packets are allowed to pass to the local machine and which of the incoming packets should be dropped since they are not targeted to the local machine. Received packets that are targeted for the local machine can be destined to the host, to a manageability controller, or to both. This section describes how host filtering is done, and the interaction with management filtering.

See Figure 7-2. Host filtering is done in three stages:

1. Packets are filtered by L2 filters (Ethernet MAC address, unicast/multicast/broadcast). See Section 7.1.1.1.
2. Packets are filtered by VLAN if a VLAN tag is present. See Section 7.1.1.2.
3. Packets are filtered by the manageability filters (port, IP, flex, other). See Section 10.3.

A packet is not forwarded to the host if any of the following occurs:

- The packet does not pass L2 filters, as described in Section 7.1.1.1.
- The packet does not pass VLAN filtering, as described in Section 7.1.1.2.
- The packet passes manageability filtering and the manageability filters determine that the packet should not pass to the host as well (see MANC2H register).

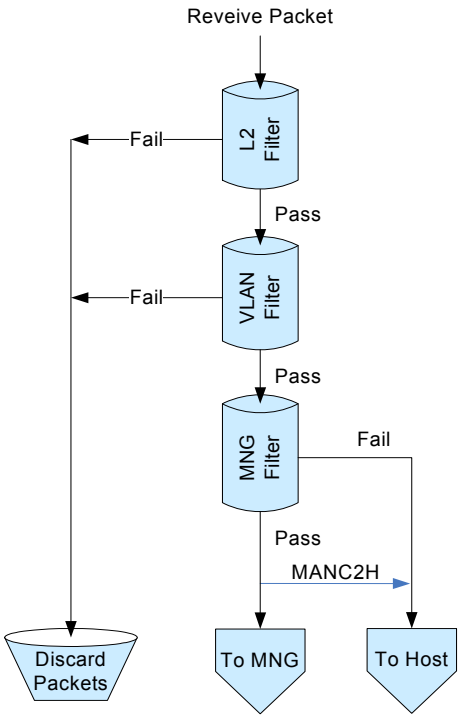


Figure 7-2 Rx Filtering Flow Chart

7.1.1.1 L2 Filtering

A packet passes successfully through L2 Ethernet MAC address filtering if any of the following conditions are met:

- Unicast packet filtering — Promiscuous unicast filtering is enabled (FCTRL.UPE=1b) or the packet passes unicast MAC filters (host or manageability).
- Multicast packet filtering — Promiscuous multicast filtering is enabled by either the host or manageability (FCTRL.MPE=1b or MANC.MCST_PASS_L2 =1b) or the packet matches one of the multicast filters.
- Broadcast packet filtering to host — Promiscuous multicast filtering is enabled (FCTRL.MPE=1b) or Broadcast Accept Mode is enabled (FCTRL.BAM = 1b).
- Broadcast packet filtering to manageability — Always enabled depending on the MDEF filters.

7.1.1.1.1 Unicast Filter

The Ethernet MAC address is checked against the 128 host unicast addresses, 4 KB hash-based unicast address filters and four management unicast addresses (if enabled). The host unicast addresses are controlled by the host interface (the manageability controller must not change them). The other four addresses are dedicated to management functions and are only accessed by the manageability. The destination address of an incoming packet must exactly match one of the pre-configured host address filters or the manageability address filters. These addresses can be unicast or multicast. Those filters are configured through Receive Address Low (RAL), Receive Address High (RAH), Manageability Ethernet MAC Address Low (MMAL) and Manageability Ethernet MAC Address High (MMAH) registers. In addition, there are 4 KB unicast hash filters used for host defined by the PFUTA registers. The unicast hash filters are useful mainly for virtualization settings in those cases that more than 128 filters might be required.

Promiscuous Unicast — Receive all unicasts. Promiscuous unicast mode can be set/cleared only through the host interface (not by the manageability controller) and it is usually used when the LAN device is used as a sniffer.

7.1.1.1.2 Multicast Filter (Partial)

The 12-bit portion of the incoming packet multicast address must exactly match the multicast filter address in order to pass multicast filter. These bits (out of 48 bits of the destination address) can be selected by the *MO* field in the MCSTCTRL register. The entries can be configured only by the host interface and cannot be controlled by the manageability controller.

Promiscuous Multicast — Receive all multicasts. Promiscuous multicast mode can be set/cleared only through the host interface (not by the manageability controller) and it is usually used when the LAN device is used as a sniffer.

7.1.1.2 VLAN Filtering

The 82599 provides exact VLAN filtering for host traffic and manageability traffic, as follows:

- Host VLAN filters are programmed by the VFTA[n] registers.
- Manageability VLAN filters are activated by the MDEF filters. One of eight VLAN tags are programmed by the MAVTV[7:0] registers while enabled by the MFVAL register.
- A VLAN match might relate to the *CFI* bit in the VLAN header. It is enabled for host filtering only by the VLNCTRL.CFIEN while the expected value is defined by the VLNCTRL.CFI.

If double VLAN is enabled (see Section 7.4.5), filtering is done on the second (internal) VLAN tag. All the filtering functions of the 82599 ignore the first (external) VLAN in this mode.

Receive packet that passes L2 layer filtering successfully is subjected to VLAN header filtering as illustrated in Figure 7-3:

1. If the packet does not have a VLAN header, it passes to the next filtering stage.
2. Else, if the packet is broadcast and MANC.RCV_TCO_EN bit is set, then it passes to the next filtering stage.

3. Else, if the packet passes a valid manageability VLAN filter and at least one VLAN_AND bit is set in the MDEF[n] registers, then it passes to the next filtering stage.
4. Else, if host VLAN filters are not enabled (VLNCTRL.VFE = 0b), the packet is forwarded to the next filtering stage.
5. Else, if the packet matches an enabled host VLAN filter and CFI checking (if enabled), the packet is forwarded to the next filtering stage.
6. Else, if manageability VLAN filtering is not required (MANC.Bypass_VLAN is set), the packet is forwarded to the next filtering stage as a potential candidate only for manageability.
7. Otherwise, the packet is dropped.

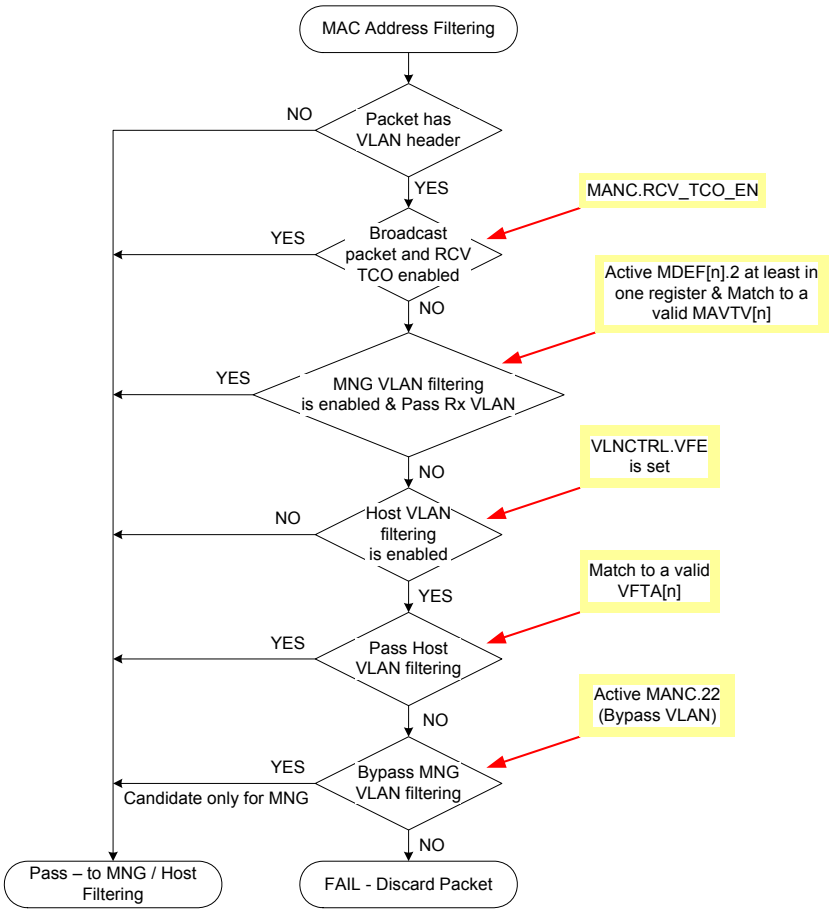


Figure 7-3 VLAN Filtering

7.1.1.3 Manageability / Host Filtering

Packets that pass the MAC address filters and VLAN address filters described in the previous sections are subjected to MNG / Host filtering shown in Figure 7-4. The Manageability filters are described in Section 10.3. Packets that are not accepted for Manageability become automatically candidates for the host queue filters described in Section 7.1.2. Packets that pass the Manageability filters may still be posted to the host as well if they match the BMC to host filters defined by the MANC2H register.

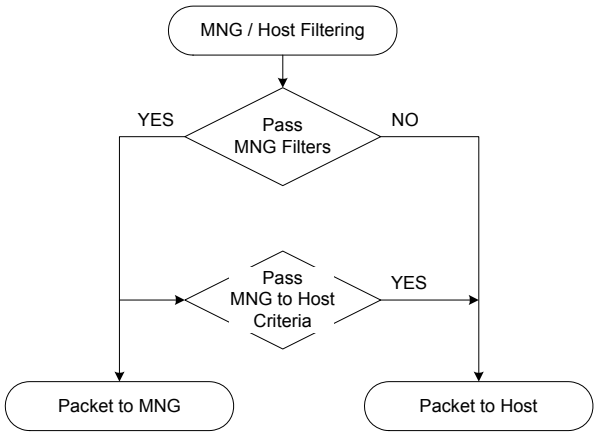
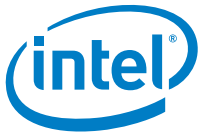


Figure 7-4 Manageability / Host Filtering

7.1.2 Rx Queues Assignment

The following filters/mechanisms determine the destination of a received packet. These filters are described briefly while more detailed descriptions are provided in the following sections:

- Virtualization — In a virtual environment, DMA resources are shared between more than one software entity (operating system and/or device driver). This is done by allocating receive descriptor queues to virtual partitions (VMM, IOVM, VMs, or VFs). Allocating queues to virtual partitions is done in sets, each with the same number of queues, called queue pools, or pools. Virtualization assigns to each received packet one or more pool indices. Packets are routed to a pool based on their pool index and other considerations such as DCB and RSS. See Section 7.1.2.2 for more on routing for virtualization.
- DCB — DCB provides QoS through priority queues, priority flow control, and congestion management. Packets are classified into one of several (up to eight) Traffic Classes (TCs). Each TC is associated with a single unique packet buffer. Packets that reside in a specific packet buffer are then routed to one of a set of Rx queues based on their TC value and other considerations such as RSS and virtualization. See Section 7.7 for details on DCB.
  - DCB is enabled via the *RT Enable* bit



- Receive Side Scaling (RSS) — RSS distributes packet processing between several processor cores by assigning packets into different descriptor queues. RSS assigns to each received packet an RSS index. Packets are routed to one of a set of Rx queues based on their RSS index and other considerations such as DCB and virtualization. See Section 7.1.2.8 for details.
- L2 Ethertype Filters — These filters identify packets by their L2 Ethertype and assigns them to receive queues. Examples of possible uses are LLDP packets, and 802.1X packets. See Section 7.1.2.3 for details. The 82599 incorporates eight Ethertype filters.
- FCoE Redirection Table — FCoE packets that match the L2 filters might be directed to a single legacy Rx queue or multiple queues to ease multi-core processing. See Section 7.1.2.4 for details. See also Section 7.13.3.3 for Large FC receive and direct data placement.
- L3/L4 5-tuple Filters — These filters identify specific L3/L4 flows or sets of L3/L4 flows. Each filter consists of a 5-tuple (protocol, source and destination IP addresses, source and destination TCP/UDP port) and routes packets into one of the Rx queues. The 82599 incorporates 128 such filters. See Section 7.1.2.5 for details.
- Flow Director Filters — These filters are an expansion of the L3/L4 5-tuple filters that provides up to additional 32 K filters. See Section 7.1.2.7 for details.
- TCP SYN Filters — might route TCP packets with their SYN flag set into a separate queue. SYN packets are often used in SYN attacks to load the system with numerous requests for new connections. By filtering such packets to a separate queue, security software can monitor and act on SYN attacks. See Section 7.1.2.6 for details.

A received packet is allocated to a queue based on the above criteria and the following order:

- Queue by L2 Ethertype filters (if match)
- Queue by FCoE redirection table (relevant for FCoE packets)
- If SYNQF.SYNQFP is zero, then
  - Queue by L3/L4 5-tuple filters (if match)
  - Queue by SYN filter (if match)
- If SYNQF.SYNQFP is one, then
  - Queue by SYN filter (if match)
  - Queue by L3/L4 5-tuple filters (if match)
- Queue by flow director filters
- Define a pool (in case of virtualization)
- Queue by DCB and/or RSS as described in Section 7.1.2.1 and Section 7.1.2.2.

7.1.2.1 Queuing in a Non-virtualized Environment

Table 7-1 lists the queuing schemes. Table 7-2 lists the queue indexing. Selecting a scheme is done via the *Multiple Receive Queues Enable* field in the MRQ register.

Table 7-1 Rx Queuing Schemes Supported (No Virtualization)

DCB	RSS	DCB / RSS Queues	Special Filters[1]
No	No	1 queue Rx queue 0	Supported
No	Yes	16 RSS queues	Supported
Yes	No	8 TCs x 1 queue 4 TCs x 1 queue RSS assign Rx queue 0 of each TC	Supported
Yes	Yes	8 TCs x 16 RSS 4 TCs x 16 RSS	Supported

1. Special filters include: L2 filters, FCoE redirection, SYN filter and L3/L4 5-tuple filters. When possible, it is recommended to assign Rx queues not used by DCB/RSS queues.

Table 7-2 Queue Indexing Illustration in Non-virtualization Mode

Queue Index bits	6	5	4	3	2	1	0
RSS	0	0	0	RSS			
DCB(4) + RSS	TC		0	RSS			
DCB(8) + RSS	TC			RSS			

A received packet is assigned to a queue according to the ordering shown in Figure 7-5):

- DCB and RSS filters and FCoE redirection — Packets that do not meet any of the filtering conditions described in Section 7.1.2 are assigned to one of 128 queues as listed in Table 7-1. The following modes are supported:
  - — No DCB, No RSS and No FCoE redirection — Queue 0 is used for all packets.
  - — RSS only — A set of 16 queues is allocated for RSS. The queue is identified through the RSS index. Note that it is possible to use a subset of these queues.
  - — DCB only — A single queue is allocated per TC to a total of eight queues (if the number of TCs is eight), or to a total of four queues (if the number of TCs is four). The queue is identified through the TC index.
  - — DCB with RSS — A packet is assigned to one of 128 queues (8 TCs x 16 RSS) or one of 64 queues (4 TCs x 16 RSS) through the DCB traffic class of the packet and the RSS index. The TC index is used as the MS bit of the Rx queue index, and the LSBits are defined by the RSS index.
  - — FCoE redirection — Up to eight queues can be allocated for FCoE traffic by the FCoE redirection table defined by FCRETA[n] registers.

When operating in conjunction with DCB, the number of RSS queues can vary per DCB TC. Each TC can be configured to a different number of RSS queues (0/1/2/4 queues). The output of the RSS redirection table is masked accordingly to generate an RSS index of the right width. When configured to less than the maximum number of queues, the respective MS bits of the RSS index are set to zero. The number of RSS queues per TC is configured in the RQTC register.

- Example — Assume a 4 TCs x 16 RSS configuration and that the number of RSS queues for TC=3 is set to 4. The queue numbers for TC=3 are 32, 33, 34, and 35 (decimal).

Figure 7-5 depicts an example of allocation of Rx queues by the various queue filters previously described for the following case:

- DCB and RSS enabled to 4 TCs x 16 RSS queues
- RSS is used at various width per TC
- SYN filter allocated
- Ethertype filters are used
- 5-tuple filters are used

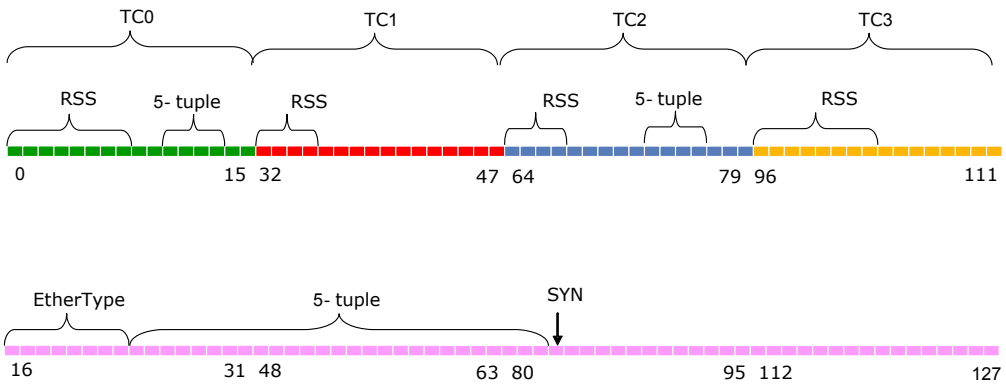


Figure 7-5 Example of Rx Queue Allocation (Non-Virtualized)

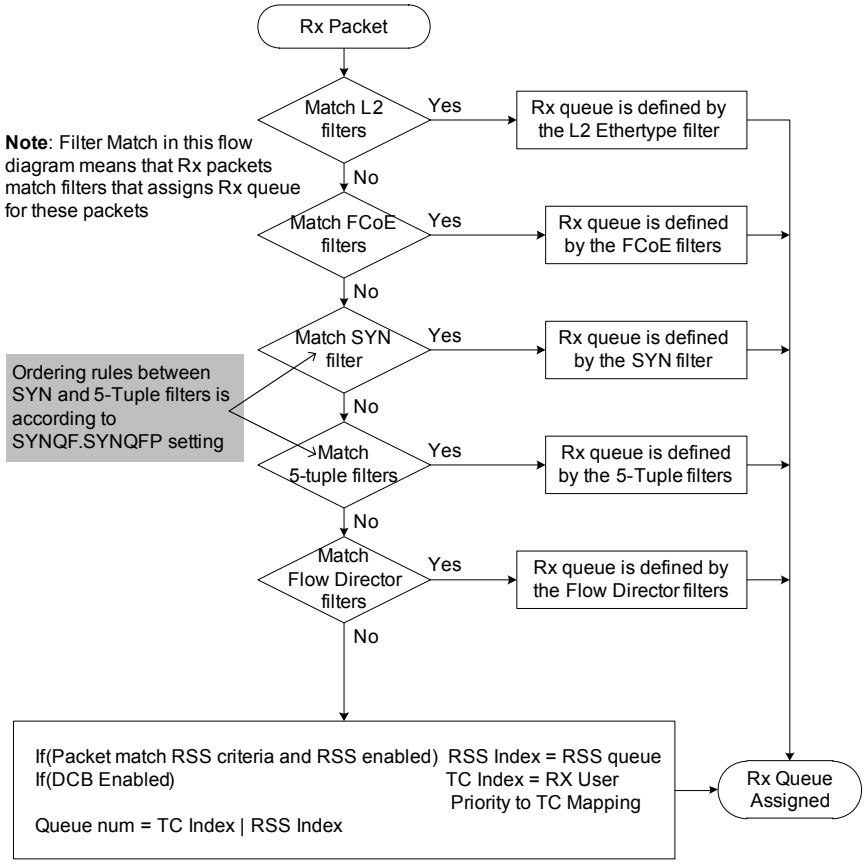


Figure 7-6 Rx Queuing Flow (Non-Virtualized)

7.1.2.2 Queuing in a Virtualized Environment

The 128 Rx queues are allocated to a pre-configured number of queue sets, called pools. In non-IOV mode, system software allocates the pools to the VMM, an IOVM, or to VMs. In IOV mode, each pool is associated with a VF.

Incoming packets are associated with pools based on their L2 characteristics as described in Section 7.10.3. This section describes the following stage, where an Rx queue is assigned to each replication of the Rx packet as determined by its pools association.

Table 7-3 lists the queuing schemes supported with virtualization. Table 7-4 lists the queue indexing.