

[9:03 AM] Ramesh Muthukrishna

but its customisable u cud use it for other purposes too

like 1

[9:04 AM] Ramesh Muthukrishna

coz its a cache

[9:04 AM] Jani Ahamed Habeeb Mohamed

ok,, also regarding this point where you told whether we need I3 processing or I2 processing..

first we will check vlan, then dmac, if dest mac is my mac, we would skip I2 processing and go for I3 lookup right?

[9:05 AM] Ramesh Muthukrishna

yes, we first do vlan membership check, then if dmac is my\_mac go to I3 pipeline directly

[9:06 AM] Jani Ahamed Habeeb Mohamed

but what if we are the actual destination the packet is meant for.. i.e dest mac matches ours

[9:07 AM] Ramesh Muthukrishna

destmac shud match our interface Ip anyways, if destip also matches my\_ip then mostly we will punt the packet to cpu coz we will have interface.ip/32 route pointing to cpu in exact match table

like 1

[9:08 AM] Ramesh Muthukrishna

only invxlan tunnelling cases even when destip matches our IP, we might not punt the packet to cpu

[9:09 AM] Jani Ahamed Habeeb Mohamed

ok, so both dest mac/dest ip match,, then we will consume the packet,, if only dst mac matched and not the dest ip,, we do I3 lookup

[9:09 AM] Ramesh Muthukrishna

destip match or not nu we confirm by doing I3 lookup only coz there is no my\_ip table in bcm

[9:10 AM] Ramesh Muthukrishna

my\_ip is programmed in exact match table. I3 l3table show la u will find

[9:10 AM] Ramesh Muthukrishna

I3 l3table show is for /32 prefix, I3 defip show for lpm taboe

[9:12 AM] Jani Ahamed Habeeb Mohamed

super, got it,, nandri!

[9:13 AM] Ramesh Muthukrishna

My\_station\_tcam is for my\_mac table

like 1

[9:13 AM] Jani Ahamed Habeeb Mohamed

ok

[9:13 AM] Ramesh Muthukrishna

can u try if this works

[9:13 AM] Jani Ahamed Habeeb Mohamed

BCM.0> list My\_station\_tcam

Memory: MY\_STATION\_TCAM.ipipe0 address 0x58240000

Flags: valid cbp cachable(on)

Blocks: ipipe0/dma/slam (1 copy, 1 dmaable, 1 slamable)

Entries: 256 with indices 0-255 (0x0-0xff), each 29 bytes 8 words

Entry mask: -1 -1 -1 -1 -1 -1 -1 0x0000001f

Description: TCAM for MPLS My Station lookups - no associated data.

L2 Ingress packet:

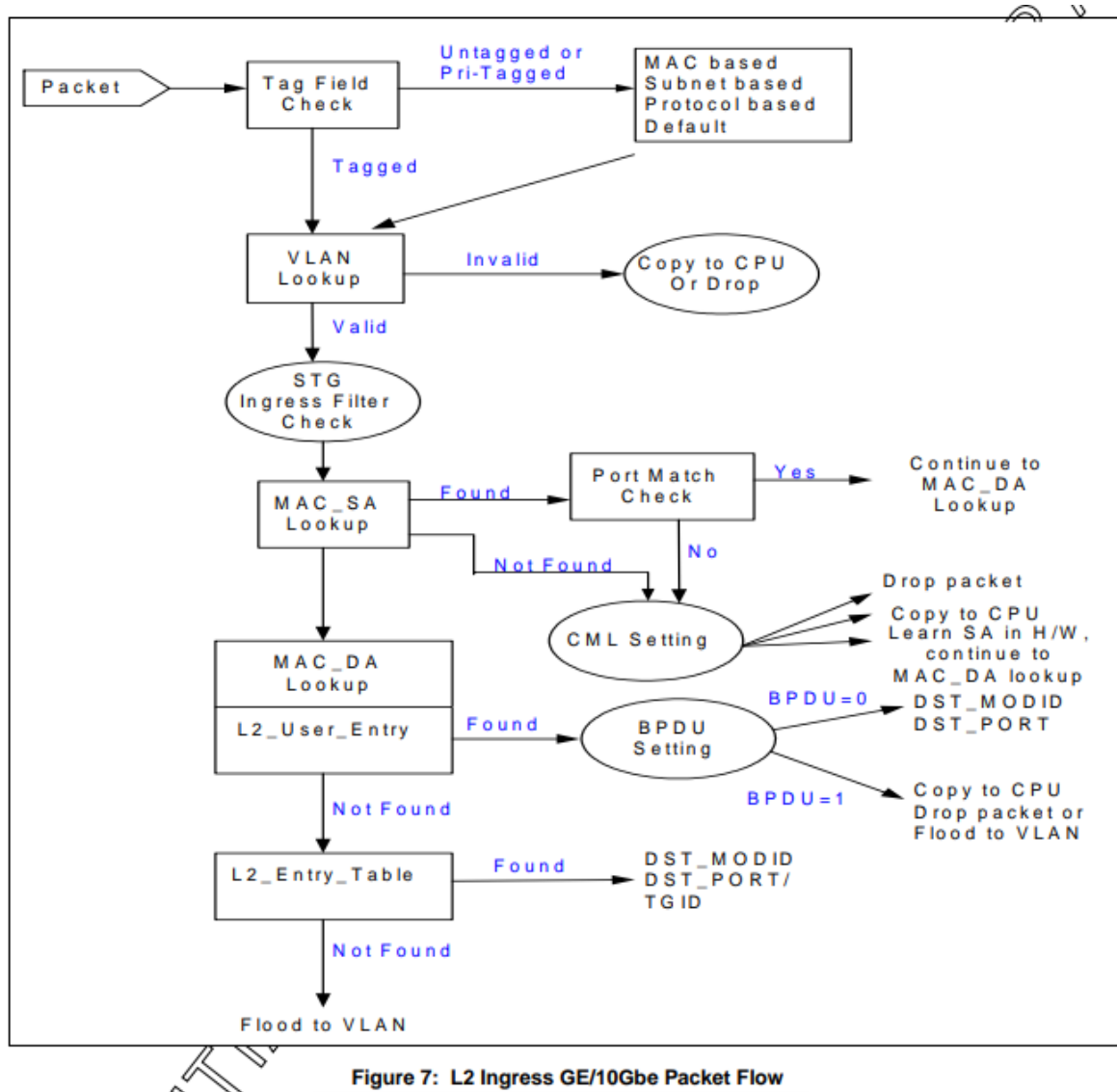


Figure 7: L2 Ingress GE/10Gbe Packet Flow

L2 unicast packet arrives: check of tagged/untagged

If untagged: vlan id is picked from MAC based/Subnet based/Protocol based/Port based vlan table

If tagged: vlan id is picked from packet itself, if vlan id invalid (EN\_IFILTER), drop, else lookup and get vlan/stp port bitmap

Packet is now at learning phase:

If SRC MAC and VLAN are present in I2 table – already learnt, otherwise based on CML setting, packet punted to CPU/learn in HW or drop

Next Dst mac lookup – If entry is found in L2 user entry table (bpdu table – immune to ageing 128 entries), then destination is based on BPDU setting.

BPDU setting 0 – packet forwarded based on dst mod id/ dst port

BPDU setting 1 – drop/copy to cpu/flood to vlan

Dst mac not found in L2 user entry but in L2 entry table(dst mac table) – then forward to dst modid / dst port, otherwise flood to vlan

L2 multicast packet : Same as L2 unicast, when it reaches the DMAC lookup in L2 entry table and matches, it is further indexed to 10-bit L2MC entry table. The L2MC table will dictate which portbit map this packet should egress out based on PFM field

PFM(port filtering mode) is vlan based, PFM 0 means flood to members of vlan, PFM 1 means forward to ports mentioned in L2MC, PFM 2 – drop

L2 broadcast packet – same as L2 unicast, but after learning, flood to all members of vlan

L3 ingress packet:

During DstMAC lookup in L2 User entry or L2 Entry table, if dst mac matches My MAC, then I3 logic kicks in

Or During DstMAC lookup in L2 user entry or L2 Entry table, if dst mac entry is there and I3\_enable bit is set, I2 logic is skipped and I3 logic kicks in

Router MAC+Vlan is programmed into L2 table with I3 enable bit set (so this means my mac matched dest mac, then I3 lookup happens)

In I3 state – first is src ip lookup (I3\_table) – no match, no hit is set and packet can be given to CPU for src ip miss and dst ip match happens next.

If src ip matches, then I3 hit bit is set, in case of station movement packet is punted to CPU and then I3 table is updated so that further dst ip lookup results in proper egress port.

During Dst Ip lookup, I3 table for directly connected hosts, I3 def ip table for Ipm matches

If match found – entry indicates dest port/next hop mac address and index into I3 interface table (which contains router mac and vlan) for egress port

For successfully routed packets , the smac, dstmac and vlan id are replaced and IP ttl field is decremented and IP checksum/ethernet fcs is recalculated

Hit bit is refreshed for ageing purposes

Higig frames are used to send control packets (CPU determines the forwarding logic – BPDU,RIP,OSPF,ARP,learning frame,CPU to CPU)

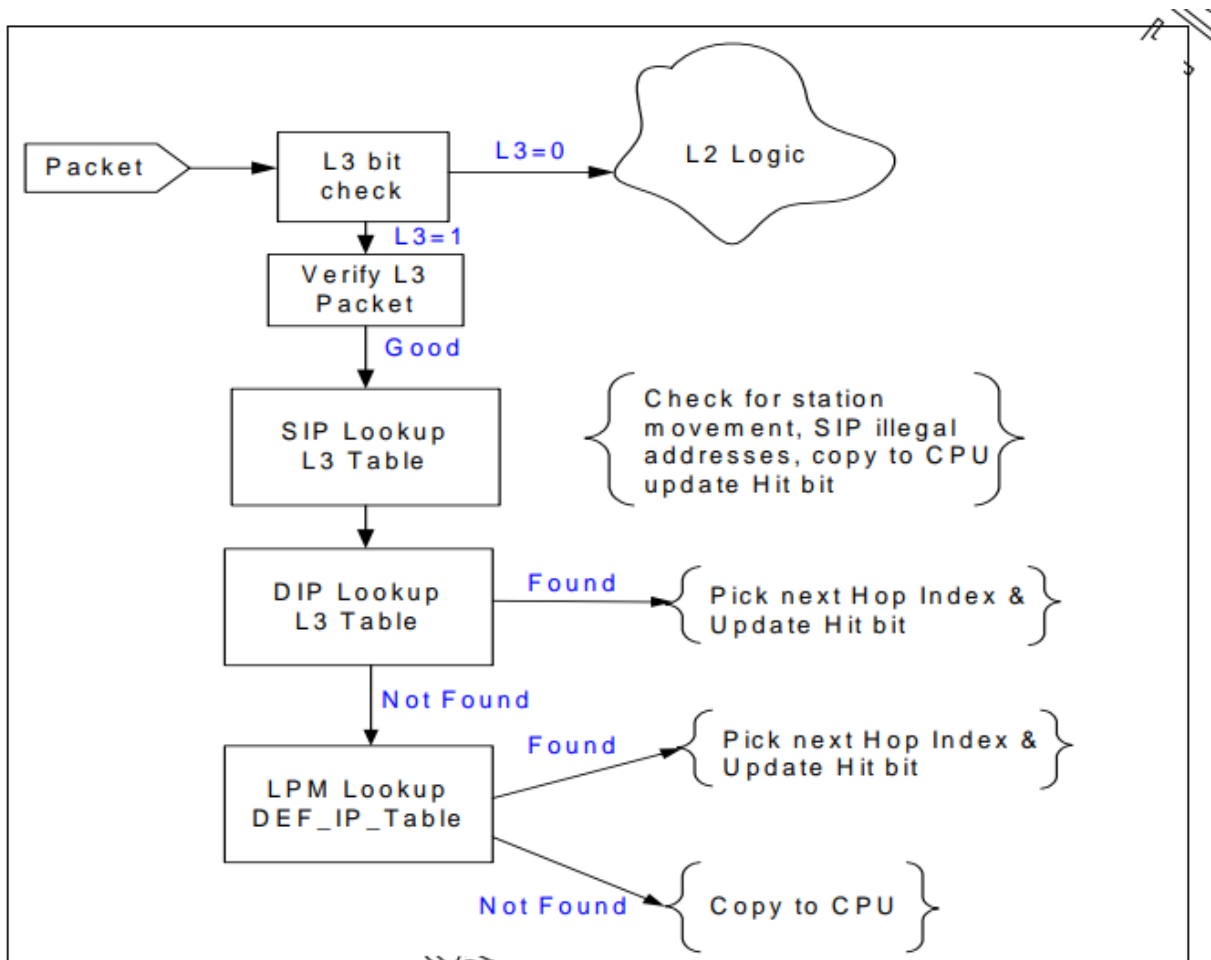


Figure 8: L3 Ingress GE/10Gbe Packet Flow

L2 table lookup is based on dst mac + vlan

Match : specified port

No match Broadcast : Flood to entire VLAN(DLF)

16K L2 hash table (dmac/vlan) – 2k buckets with 8 entries each

To avoid station movement set static entry along with hit bit

L3 table is 8K entries with 1024 buckets and 8 entries deep

L3 def ip table – 6K entries (tcam)

L3 interface table – 4K entries

After LPM from l3 def table – ECMP is done to determine the path based on CRC32 hashing of src.dstip/protocol etc

Storm control – based on bcast/mcast/dlf storm control bits – port based