BCM88800

Device Errata

Errata

Broadcom reserves the right to make changes without further notice to any products or data herein to improve reliability, function, or design. Information furnished by Broadcom is believed to be accurate and reliable. However, Broadcom does not assume any liability arising out of the application or use of this information, nor the application or use of any product or circuit described herein, neither does it convey any license under its patent rights nor the rights of others.

Table of Contents

Chapter 1: Introduction

1.1 Scope

This errata sheet lists all known errata for the Broadcom® BCM88800 (Jericho2c) self-routing switching element.

1.2 Summary Lists

1.2.1 Traffic Management Errata

Table 1: BCM88800 Traffic Management Errata Summary by Number

Functional Errata Number	Description and Reference
EID#8022	EID#8022 OTM Shapers Can Get Stuck When the Packet Size Is Larger than the Shaper Max-Burst Value
EID#8024	EID#8024 Statistic Record Can Contain the Wrong Disposition Indication
EID#8025	EID#8025 ILKN FEC Alignment Failure Indications Are Shared Between Two ILKN Interfaces in the Same ILKN Core
EID#8030	EID#8030 ILKN Junk Data Becomes Stuck in the RX High-Rate FIFO Segmentation Module after Link Down
EID#8034	EID#8034 Egress Packet Scheduler Weighted Fair Queuing Overflow
EID#8035	EID#8035 Interlaken Start-of-Packet Segment Might Be Too Short for SRv6 Packets
EID#8038	EID#8038 Multicast Packets Cannot Be Counted in the Egress Traffic Manager
EID#8041	EID#8041 Channelized Over Ethernet Flow Control
EID#8043	EID#8043 Egress TC/DP Mapping of MC Packets Can Cause Corruption of ECGM
EID#8044	EID#8044 Mirror-on-Drop for Packets Shorter than 256 Bytes

1.2.2 Packet Processing Errata

Table 2: BCM88800 Packet Processing Errata Summary by Number

Functional Errata Number	Description and Reference
EID#9004	EID#9004 Congestion Notification Indication Based on E2E Latency Measurement for Forwarding Header + 1 Is Incorrect (ECN Application)
EID#9100	EID#9100 IP Multicast Bridge Fallback Does Not Work
EID#9120	EID#9120 Some Events Generated in the MAC Table Are Not Sent to the CPU
EID#9121	EID#9121 Trap for an Unknown MAC Address Cannot Be Used
EID#9127	EID#9127 OAMP Flexible CRC Verification Cannot Coexist with OAMP LM/DM Per Priority
EID#9141	EID#9141 IP Multicast Fallback to Bridge Corrupts OutLIF1, OutLIF2, and OutLIF3 Pointers at the ERPP
EID#9144	EID#9144 InLIF Information at the iPMF Might Be Incorrect When Terminating an IP Tunnel
EID#9154	EID#9154 Events from the RX Processor Can Corrupt Protection Packets
EID#9163	EID#9163 SLR Measurement Interval Does Not Work Correctly

Chapter 2: Traffic Management Errata

This chapter describes the errata for the BCM88800 traffic management (TM) feature.

2.1 EID#8022 OTM Shapers Can Get Stuck When the Packet Size Is Larger than the Shaper Max-Burst Value

Description Egress OTM shapers have a configurable max-burst value controlled by `bcm_cosq_control_set(unit, gport, 0, bcmCosqControlBandwidthBurstMax, 16383).`

Erratum Due to an internal error, when a packet size greater than the max-burst value goes through the OTM, the egress shaper might get stuck (deadlock). No packets can flow out from this port until a software reset routine or reset is activated on the device.

Workaround Set the max-burst value higher than the maximum possible packet size. The recommended max-burst value is 16 KB.

2.2 EID#8024 Statistic Record Can Contain the Wrong Disposition Indication

Description The statistic interface supports sending a statistics (billing) record per packet. The record can contain a disposition field.

Erratum If the device is operating in single-report mode for all MC replications, the statistics record for a discarded packet with a valid SNIF copy might contain incorrect data.

In this scenario, the disposition field in the record is wrongly reported as *forwarded* instead of *discarded*.

Workaround Change the MC report mode to report per copy (SOC property `stat_if_report_multicast_single_copy = 0`).

2.3 EID#8025 ILKN FEC Alignment Failure Indications Are Shared Between Two ILKN Interfaces in the Same ILKN Core

Description

Two of the BCM88800 Interlaken (ILKN) cores connect to PAM4-capable lanes. Each of these ILKN cores can be configured as one or two interfaces:

- ILKN core 6 (interfaces 12 and 13)
- ILKN core 7 (interfaces 14 and 15)

These two cores include ILKN FEC (Forward Error Correction) encoders and decoders.

For an ILKN interface using a SerDes rate of up to 30G in NRZ mode, FEC is optional.

For an ILKN interface using a SerDes rate above 30G in PAM4 mode, FEC is mandatory.

Erratum

The erratum affects applications configuring either core 6 or core 7 as two ILKN interfaces, and when both interfaces use FEC.

The FEC alignment error indication (no alignment) is shared for the two interfaces. FEC alignment loss in one interface also affects the other interface. Traffic stops on both interfaces until software handles the FEC alignment loss of the failed interface.

NOTE:

- There are no limitations for a single interface in an ILKN core.
- There is no dependency between interfaces in different ILKN cores.
- There is no impact if one of the interfaces in the ILKN core does not use FEC.

Workaround

When configuring core 6 or core 7 as two interfaces, do the following:

- Avoid using two interfaces over PAM4 (where FEC is mandatory).
- Disable FEC for interfaces operating in NRZ mode.

2.4 EID#8030 ILKN Junk Data Becomes Stuck in the RX High-Rate FIFO Segmentation Module after Link Down

Description

The ILKN core interface is a segmented interface that consists of several segments of 32B each cycle. The 32B segmented interface arrives at the RX high-rate FIFO (HRF) segmentation module to be packed to 64B.

If a segment of 32B without an end-of-burst (EOB) indication arrives at the HRF segmentation module, it is buffered until the next word arrives.

Erratum

When an ILKN link-down event occurs, the last data before the link went down might not have an EOB indication.

The issue occurs because the data without the EOB flag remains stuck in the HRF segmentation module, and when the link comes back up, the first packet on this port is merged with the previous junk data. As a result, the first packet on this port after the link comes up might be corrupted.

Workaround

NOTE: This workaround is only partial because the HRF reset is not per port.

When the core has a single ILKN port or if both ILKN ports are down, reset the HRF segmentation module to delete the junk data from the buffer.

This workaround is not valid when the core has two ILKN ports, and only one of the ILKN ports has a link-down event. In this case, it is possible to send one CPU packet to clean the junk data from the buffer.

2.5 EID#8034 Egress Packet Scheduler Weighted Fair Queuing Overflow

Description Traffic class group (TCG) shapers are used in eight-priority and four-priority packet scheduler modes. In these modes, traffic below the TCG shaper threshold is classified as committed information rate (CIR) traffic, and traffic above the TCG shaper threshold is classified as excess information rate (EIR) traffic.

All CIR traffic enters a round-robin (RR) or strict priority (SP) arbiter.

All EIR traffic enters a weighted fair queuing (WFQ) arbiter.

After the RR/SP and WFQ arbiters, the CIR traffic has a strict higher priority over the EIR traffic.

Erratum This issue is relevant when using eight or four priorities only.

When applying a queue-pair shaper, priority flow control (PFC), or both on one of the priorities, the EIR WFQ is not honored.

Workaround Disable queue-pair shapers and disable PFC.

2.6 EID#8035 Interlaken Start-of-Packet Segment Might Be Too Short for SRv6 Packets

Description The next SRv6 segment identifier (SID) is copied from the incoming segment routing header (SRH) SID list to the packet-processing pipeline.

Erratum When using Interlaken (ILKN), the size of the first burst is not guaranteed to be 256B (it is only guaranteed to be the delta between Ilkn_burst_max and Ilkn_burst_min). Therefore, the first burst received by the ingress receive editor (IRE) must be larger than the location copied to the IRPP header to ensure the correct processing.

Workaround Working with ILKN normal mode (and not enhanced mode) on the TX side of the remote peer guarantees a first burst with a size of 256 bytes.

If the DNX device ILKN is working in full-packet mode, the ILKN RX enables burst merge (by default) to increase the size of transactions to the IRE. In this case, the first segment is up to 512B.

2.7 EID#8038 Multicast Packets Cannot Be Counted in the Egress Traffic Manager

Description The egress traffic manager (ETM) can be configured to send counting information per egress queue, including both unicast (UC) and multicast (MC) queues.

Erratum MC packets cannot be counted in the ETM.

Workaround Use the counter in the egress receive packet processor (ERPP).

Generate a statistics command with the OTM port as the statistics object ID (SOID) and the traffic class.

2.8 EID#8041 Channelized Over Ethernet Flow Control

Description The device supports Channelized Over Ethernet (COE) traffic and supports PFC or pause frames over COE packets (known as COE FC).

A COE FC frame holds a *timeout* field that indicates the time period for which traffic on the specific channel should be paused. If timeout is zero, traffic on the specific channel should be transmitted (known as Xon).

A PFC frame holds eight timeout fields, one per priority.

Erratum Due to this erratum, timers from the COE FC frames that are associated with a channel might not begin counting. This means that the COE FC application cannot rely on the timeout mechanism to block the traffic of a specific channel (and priority) for the correct period of time.

Note that the erratum is only affecting the timeout value, and the pause state (Xoff or Xon) is applied correctly.

Workaround The COE FC application should apply the Xon/Xoff approach.

When the traffic on a specific channel (and maybe priority) should be paused (Xoff), the receiver port should transmit a COE FC frame with the timeout value set to the maximum value (0xFFFF).

When the traffic on a specific channel (and maybe priority) should be released (Xon), the receiver port should transmit a COE FC frame with the timeout value set to zero (0x0).

2.9 EID#8043 Egress TC/DP Mapping of MC Packets Can Cause Corruption of ECGM

Description A MC packet is mapped in the EGQ to either high or low service pools (SP) and one of eight MC-TC, based on its TC and DP fields.

Erratum Due to this erratum, the SP and MC-TC mapping is wrong. This might result in wrong counter updates in the ECGM, status corruption, up to a device reset.

Workaround Work with a single SP and a single MC-TC in the ECGM. In this case, all the MC traffic is mapped to the same SP and the same MC-TC counters in the ECGM.

2.10 EID#8044 Mirror-on-Drop for Packets Shorter than 256 Bytes

Description The device supports the Mirror-On-Drop feature (also known as dropped-packet reports). This feature allows sending to an external analyzer a copy of a packet that was eligible for being dropped. This mirroring might be set per drop reason and scope (see the "Mirror-on-Drop" section in *Traffic Management Architecture* [88800-DG1xx]).

The feature allows sending a copy of the full packet or cropping the packet and only sending the first 256 bytes of the packet.

Erratum Due to this erratum, if a packet that is shorter than 256 bytes is mirrored by the Mirror-On-Drop mechanism, and the mirror profile is set to crop the packet's copy, when the cropped copy is moved to the DRAM (S2D), the VOQ DRAM size is corrupted. As a result, the DRAM size of the queue is never 0, and the queue is never empty.

Workaround A VOQ that is used for cropped Mirror-On-Drop packets should be set as OCB.

Chapter 3: Packet Processing Errata

3.1 EID#9004 Congestion Notification Indication Based on E2E Latency Measurement for Forwarding Header + 1 Is Incorrect (ECN Application)

Description

Congestion Notification Information (CNI) is a standard allowing the congestion level to be signaled on IPvX headers. This is used for ECN applications.

In the device, there are four options to calculate the CNI:

- Based on VOQ threshold
- Based on ingress queuing latency measurement (ingress latency)
- Based on egress TM phantom queue thresholds
- ETPP latency (E2E latency)

It is possible to signal the CNI level on the forwarding header or on the forwarding header + 1 (the header above the forwarding header) if the forwarding header is Ethernet.

Updating the CNI value over the IPvX header is done at the ETPP termination stage.

Erratum

When the CNI is calculated using E2E latency, and the header to update is the forwarding + 1 IPvX header, the CNI value is not correct.

Workaround

Use ingress latency to calculate the CNI value for this use case.

3.2 EID#9100 IP Multicast Bridge Fallback Does Not Work

Description

In some use cases, IPvX compatible multicast (MC) routed packets that do not find a destination are required to perform IP multicast bridge fallback.

IP multicast bridge fallback means the packet is treated as a bridged packet with an unknown destination.

To perform the IP multicast bridge fallback, the IRPP forwarding performs the following procedure:

1. Identifies that the packet is IP multicast bridge fallback eligible.
2. Returns to the ETH header parameters.

   IP multicast bridge fallback is performed only for IPvX(MC)oETHo[Tunnels], that is, an Ethernet layer below the IP layer is guaranteed, and there might be tunnel headers below the Ethernet layer.

   – Decreases the forwarding header offset by one, so that the forwarding header offset will point to the Ethernet layer.
   – Changes the forwarding-domain to the VSI of the Ethernet layer.
   – Pops the `In_LIF` and `In_LIF_Profile` stacks so that the `L2_LIF` is used (and not the `ETH_RIF`).
3. Executes the Ethernet default procedure.

Erratum

Due to this erratum, IP multicast bridge fallback does not work.

Workaround

Create a two-pass processing for IP multicast bridge fallback:

1. First pass.

   Identify the case in PMF1/2, and recycle the packet without any change (if tunnel headers reside below the Ethernet layer header, terminate the tunnel headers here).
2. Second pass.

   Set the `PTC-Profile` of the `Recycle_Port` to prevent the *Routing-Enabler* procedure from terminating the Ethernet layer.

NOTE:

- This workaround is not implemented in the SDK.
- Implement this workaround at the application level.

3.3 EID#9120 Some Events Generated in the MAC Table Are Not Sent to the CPU

Description

The MAC table (MACT) block is responsible for managing the MAC address database.

Different events can be sent from the MACT to the CPU using a DMA FIFO. The following list describes the MACT event types:

- Exceed limit – When the number of entries at the MAC table exceeds the MACT global limit or the per-VSI limit, learning does not occur.
- Delete non exist – A delete command arrives from the CPU, but the entry to remove is not found.
- Learn over stronger – A learn command reaches a MACT entry with a strength higher than the strength of the learn command, so learning does not occur.
- Transplant over stronger – A transplant command arrives from the CPU with a strength lower than the entry strength, so learning does not occur.
- Refresh over stronger – A DSP packet is received with learn information, and the strength is lower than the current MAC entry strength, so learning does not occur.
- Flush drop – The CPU initiated a rule in the flush machine, and the rule is active. A new address to learn is checked against the active rules. If a match occurs and the rule causes the address to be dropped, the CPU should be notified. Learning does not occur.

Erratum

Due to this erratum, the following events cannot be sent to the CPU:

- Delete non exist
- Learn over stronger
- Transplant over stronger
- Refresh over stronger

The CPU does not receive any indication if these events occur.

Workaround

Each event has a counter that indicates how many events have occurred, but no event details are included.

The counter is primarily useful for *Delete non exist* event types.

For details, refer to the register MACT_MACT_ERROR_DELETE_NON_EXIST_COUNTER.

3.4 EID#9121 Trap for an Unknown MAC Address Cannot Be Used

Description

When a MAC DA lookup is performed, the DA might not have a match. This type of packet is marked as an unknown DA. If this occurs, a programmable destination is selected for the packet. In many cases, the packet is flooded across the VLAN.

The MAC table lookup occurs at the FLP (forwarding) stage.

Erratum

Due to this erratum, when an unknown DA packet destination is encoded as a trap, this trap is also selected for known DA packets (where the MAC table lookup was successful).

Workaround

For unknown DA packets, use a regular destination that is not a trap.

3.5 EID#9127 OAMP Flexible CRC Verification Cannot Coexist with OAMP LM/DM Per Priority

Description The OAMP supports loss measurement and delay measurement per endpoint. If LM/DM is required per priority, the OAM-ID arriving to the OAMP can be remapped so it points to an entry that handles LM/DM for a specific priority.

In addition, the OAMP supports flexible CRC verification per endpoint to validate expected data for non-standard OAM packets. An example is 48B MAID verification for received CCMs.

Erratum Due to this erratum, flexible CRC verification cannot coexist with OAM LM/DM per priority.

Workaround Instead of using the OAM-ID and priority at the OAMP, it is possible to perform the remapping at the PMF stage. The PMF maps the OAM-ID to a new OAM-ID per priority before sending the packet to OAMP, so only the flexible CRC feature is used at the OAMP.

3.6 EID#9141 IP Multicast Fallback to Bridge Corrupts OutLIF1, OutLIF2, and OutLIF3 Pointers at the ERPP

Description The device supports IP multicast (MC) handling. When the ingress interface and egress interface are the same (InLIF = OutLIF) and the packet is routed IP MC, IP-layer handling (routing) should be canceled so that bridging is performed instead. In this case, the IP header is not changed. This functionality is called *fallback to bridge*.

Fallback to bridge is performed in the egress device in the ERPP pipe and is performed again in the ETPP pipe.

When fallback to bridge occurs, OutLIF values must be set as follows:

- OutLIF0 = OutLIF1
- OutLIF1 = OutLIF2
- OutLIF2 = OutLIF3
- OutLIF3 = NULL

Erratum Due to this erratum, when fallback to bridge is performed in the ERPP, only the OutLIF0 = OutLIF1 value is set. OutLIF1, OutLIF2, and OutLIF3 are not updated. When using ePMF, the OutLIF1, OutLIF2, and OutLIF3 qualifiers are not correct.

Workaround None.

3.7 EID#9144 InLIF Information at the iPMF Might Be Incorrect When Terminating an IP Tunnel

Description Packets arriving to the device are assigned an AC LIF. A new LIF is assigned upon successful tunnel termination.

In the BCM88800 device, the last two InLIFs are saved and might be used later for ACL rules. For example, for an IPv6oIPv4_GREoETH packet forwarded according to the IPv6 header, the following LIF stack is created:

- InLIF[0]: IPv4_GRE
- InLIF[1]: ETH_RIF
- InLIF[2]: AC_LIF

Erratum Due to this erratum, an InLIF that is MP or P2P might be pushed into the stack twice, causing the pipeline to lose InLIF[1], and might cause incorrect handling of this packet by the ACL rules.

For the example in the description, the issue causes the following InLIF stack to be created:

- InLIF[0]: IPv4_GRE
- InLIF[1]: IPv4_GRE
- InLIF[2]: ETH_RIF
- InLIF[3]: AC_LIF

Workaround None.

3.8 EID#9154 Events from the RX Processor Can Corrupt Protection Packets

Description OAMP can generate protection packets, which carry events generated by either the RMEP scanner (such as LOC, ALOC, and so on) or the RX processor (for example, RDI set/clear). These events are padded to 8B in the packet.

Erratum Due to this erratum, when the RX processor creates an event, it might place the first half (first 4B) of the event in one packet (at its end or tail) and the second half at the next packet (at its start or head).

Broken events will be lost unless reconstructed by the application.

Workaround The application must monitor the received packets and reconstruct the broken event out of the next packet.

3.9 EID#9163 SLR Measurement Interval Does Not Work Correctly

Description Synthetic loss measurement, as defined in the ITU-T Y.1731 standard, is supported in the OAMP. To support the *measurement period* defined in the standard, the OAMP has the following mechanism:

- One of eight timers (each configured to its own wraparound value) is selected per session.
- When the timer associated with the session expires, the `MeasureNextReceivedSlr` bit is set for that session.
- The loss management accounting is done when an SLR packet is received, and the `MeasureNextReceivedSlr` bit is set for the corresponding session. If the bit is not set, the counters are tracked without measurement.

Erratum For self-contained entries, the wrong bit is updated instead of the `MeasureNextReceivedSlr` bit. As a result, the entry is corrupted. Therefore, this feature must not be enabled for self-contained entries. For offloaded entries, the mechanism works, but the initial values of the counters are set to the maximum value and not to the configurable wraparound value. Therefore, after reset, a 2-minute wait period must pass before SLM statistics can be logged.

As a consequence:

- Cannot use self-contained entries with SLR
- Must wait 2 minutes before logging SLM stats with offloaded entries

Workaround There is no workaround.

BROADCOM®