

NAME

sfxge - Solarflare 10Gb Ethernet adapter driver

SYNOPSIS

To compile this driver into the kernel, place the following lines in your kernel configuration file:

```
device sfxge
```

To load the driver as a module at boot time, place the following line in loader.conf(5):

```
sfxge_load="YES"
```

DESCRIPTION

The **sfxge** driver provides support for 10Gb Ethernet adapters based on Solarflare SFC9000 and XtremeScale X2 family controllers. The driver supports jumbo frames, transmit/receive checksum offload, TCP Segmentation Offload (TSO), Large Receive Offload (LRO), VLAN checksum offload, VLAN TSO, and Receive Side Scaling (RSS) using MSI-X interrupts.

The driver allocates 1 receive queue, transmit queue, event queue and IRQ per CPU up to a maximum of 64. IRQ affinities should be spread out using cpuset(1). Interrupt moderation may be controlled through the sysctl *dev.sfxge.%d.int_mod* (units are microseconds).

For more information on configuring this device, see ifconfig(8).

A large number of MAC, PHY and data path statistics are available under the sysctl *dev.sfxge.%d.stats*. The adapter's VPD fields including its serial number are available under the sysctl *dev.sfxge.%d.vpd*.

HARDWARE

The **sfxge** driver supports all 10Gb Ethernet adapters based on Solarflare SFC9000 family controllers.

LOADER TUNABLES

Tunables can be set at the loader(8) prompt before booting the kernel or stored in loader.conf(5). Actual values can be obtained using sysctl(8).

hw.sfxge.rx_ring

The maximum number of descriptors in a receive queue ring. Supported values are: 512, 1024, 2048 and 4096.

hw.sfxge.tx_ring

The maximum number of descriptors in a transmit queue ring. Supported values are: 512, 1024,

2048 and 4096.

hw.sfxge.tx_dpl_get_max

The maximum length of the deferred packet "get-list" for queued transmit packets (TCP and non-TCP), used only if the transmit queue lock can be acquired. If a packet is dropped, the *tx_get_overflow* counter is incremented and the local sender receives ENOBUFS. The value must be greater than 0.

hw.sfxge.tx_dpl_get_non_tcp_max

The maximum number of non-TCP packets in the deferred packet "get-list", used only if the transmit queue lock can be acquired. If a packet is dropped, the *tx_get_non_tcp_overflow* counter is incremented and the local sender receives ENOBUFS. The value must be greater than 0.

hw.sfxge.tx_dpl_put_max

The maximum length of the deferred packet "put-list" for queued transmit packets, used if the transmit queue lock cannot be acquired. If a packet is dropped, the *tx_put_overflow* counter is incremented and the local sender receives ENOBUFS. The value must be greater than or equal to 0.

hw.sfxge.tso_fw_assisted

Bitmask to enable/disable usage of FW-assisted TSO version if supported by NIC firmware. FATSOv1 (bit 0) and FATSOv2 (bit 1) are supported. All enabled by default.

hw.sfxge.N.max_rss_channels

The maximum number of allocated RSS channels for the Nth adapter. If set to 0 or unset, the number of channels is determined by the number of CPU cores.

hw.sfxge.lro.table_size

Size of the LRO hash table. Must be a power of 2. A larger table means we can accelerate a larger number of streams.

hw.sfxge.lro.chain_max

The maximum length of a hash chain. If chains get too long then the lookup time increases and may exceed the benefit of LRO.

hw.sfxge.lro.idle_ticks

The maximum time (in ticks) that a connection can be idle before it's LRO state is discarded.

hw.sfxge.lro.slow_start_packets

Number of packets with payload that must arrive in-order before a connection is eligible for LRO. The idea is we should avoid coalescing segments when the sender is in slow-start because reducing the ACK rate can damage performance.

hw.sfxge.lro.loss_packets

Number of packets with payload that must arrive in-order following loss before a connection is eligible for LRO. The idea is we should avoid coalescing segments when the sender is recovering from loss, because reducing the ACK rate can damage performance.

hw.sfxge.mcdi_logging

Enable logging of MCDI protocol messages (only available if enabled at compile-time).

hw.sfxge.N.mcdi_logging

Enable or disable logging of MCDI protocol messages on a per-port basis. The default for each port will be the value of *hw.sfxge.mcdi_logging*. The logging may also be enabled or disabled after the driver is loaded using the `sysctl dev.sfxge.%d.mcdi_logging`.

hw.sfxge.stats_update_period_ms

Period in milliseconds to refresh interface statistics from hardware. The accepted range is 0 to 65535, the default is 1000 (1 second). Use zero value to disable periodic statistics update. Supported on SFN8xxx series adapters with firmware v6.2.1.1033 and later and SFN5xxx, SFN6xxx and XtremeScale X2xxx series adapters. SFN7xxx series adapters and sfN8xxx series with earlier firmware use a fixed 1000 milliseconds statistics update period. The period may also be changed after the driver is loaded using the `sysctl dev.sfxge.%d.stats_update_period_ms`.

SUPPORT

For general information and support, go to the Solarflare support website at:
<https://support.solarflare.com>.

SEE ALSO

`cpuset(1)`, `arp(4)`, `netintro(4)`, `ng_ether(4)`, `vlan(4)`, `ifconfig(8)`

AUTHORS

The **sfxge** driver was written by Philip Paeps and Solarflare Communications, Inc.