

Buffer Management

NOTE:

On Trident and Trident+ based switches, known unicast packets can be assigned to a specific queue between queues from 0 to 7, while unknown unicast packets, multicast packets, and broadcast packets can be assigned to a specific queue between queues from 0 to 3.

Introduction

When the network congestion, the port don't send data immediately, to prevent data loss, device can transmit the temporary storage of data to the data buffer.

1. Accounting Resources

The chip has the fixed-sized cells. Each cell is 208 bytes long. The first cell used by a packet contains two payloads. The first payload is a 64-byte packet descriptor, and the second payload contains 144 bytes of packet data. Every cell, after the first one used by a packet, uses the complete 208 bytes for packet data. A packet use at least one cell. The length of packet decides how many cell the packet uses. For example untagged packets 64 bytes to 144 bytes long use one cell. A 352-byte packet would require two cells. And packets from 1393-bytes to 1600-bytes long use eight cells.

2. Memory Spaces

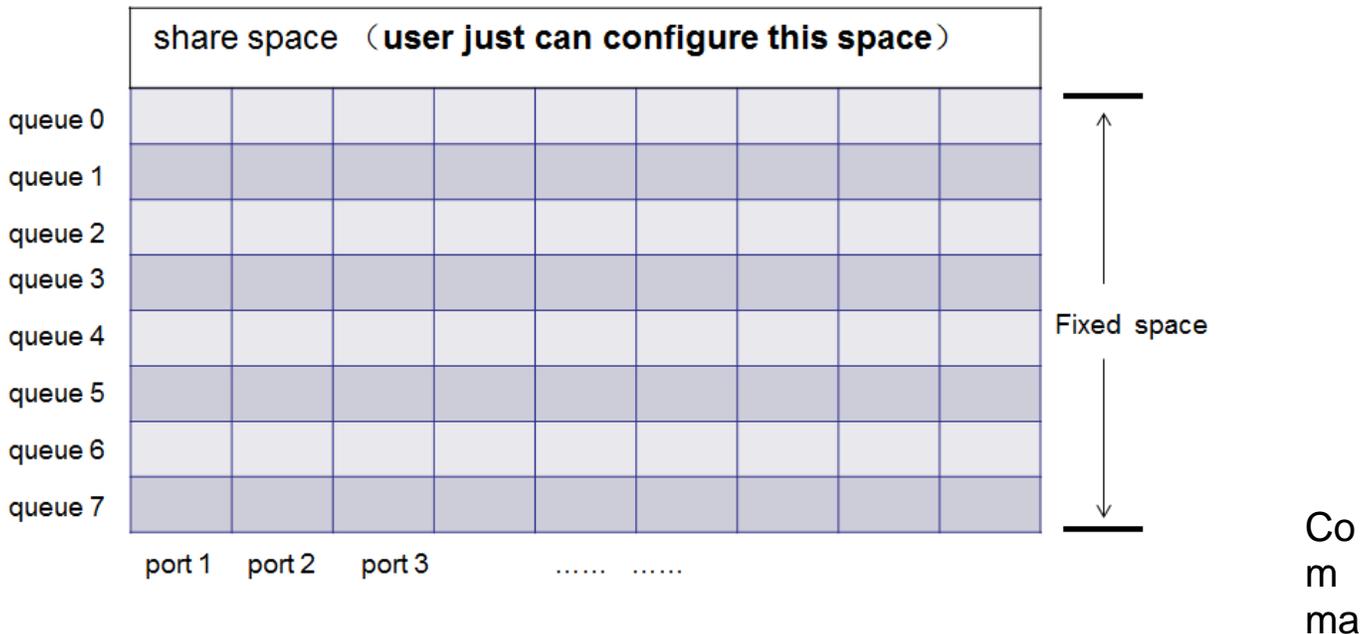
The cell data buffer is divided into fixed and shared spaces.

- Fixed space

Fixed space provide minimum guaranteed space on a per priority group and per port basis. This space cannot be shared with other ports. When the port congestion occurs, it will occupy fixed space first, shared space will be used when the fixed space is full. The Fixed space should not be too big, otherwise it will waste space resources.

- Shared space

The shared space will be used by all ports and priority groups when the fixed space run out. When the shared space is used up, the packet is discarded. If a queue don't store packet, the others queue will grab this queue space. For a queue, all of ports transmit packets will first come first served, if the used resource reach the threshold, packets will be dropped after arrived. The shared space can be configured dynamic or static threshold at the logical level of the queue.



and configuration

The device provide separate UC (unicast) and MC (multicast) queue. For the unicast queue and multicast queue. The UC queues are used to store known unicast packets and can be flexibly assigned while the MC queues are used to store DLF, multicast, broadcast and mirrored packets. While the CPU port does not support separate UC and MC queues, the CPU port does have a total of 48 queues available, and these queues are eligible for both UC and MC traffic based on CPU Cos mapping. We have different configuration to deal with share space

1. multicast queue

static mode: This is the default configuration

- The value of the static threshold is configured for the multicast queue in the shared space. This avoids the case that the port enabled flow control or PFC affects the forwarding of other ports. But the disadvantage is less efficient use of memory.

dynamic mode : Need to configure

- The memory threshold is adjusted dynamically according to the available shared memory in this mode. The thresholds will reduce when the remaining shared memory reduces. The thresholds will rise up when the remaining shared memory rises up.

For example, enable queue 0 dynamic-shared

```
root@XorPlus# set interface ethernet-switching-options buffer egress-queue 0 mc-queue-dynamic-shared true
root@XorPlus# commit
```

2. unicast queue

dynamic mode

- dynamic threshold is the default mode for the unicast queue and it can't configure static threshold.

3. dynamic threshold ratio configuration

The dynamic mode both of multicast (must enable dynamic threshold) and unicast queue can configure the max dynamic threshold ratio. The threshold is still dynamically adjusted until reaching the max ratio configured.

For example, configure dynamic threshold ratio 50% for queue 0

```
root@XorPlus# set interface ethernet-switching-options buffer egress-queue 0 shared-ratio 50
root@XorPlus# commit
```

Check the configuration result

```
root@XorPlus# run show interface egress-buffer
Cosq Buffer-Shared-Ratio Multicast-Queue-Dnamic-Shared-Buffer
0 50 Enable
1 33 Disable
2 33 Disable
3 33 Disable
4 33 Disable
5 33 Disable
6 33 Disable
7 33 Disable
```

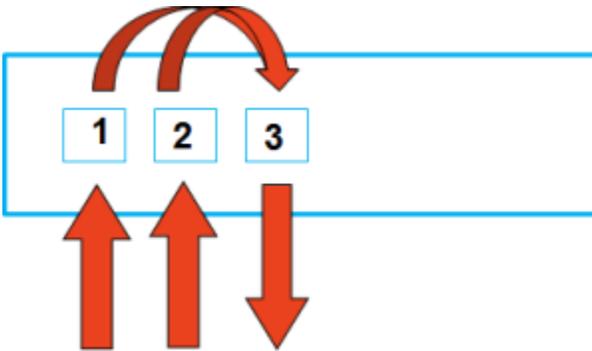
For the maximum shared-area ratio for a queue, the percentage values 0 to 100 are divided into 10 ranges. The table below shows the effective values that correspond to the configured values of ratio-value.

Value of ratio-value	Effective value
0 to 1	1
2 to 3	3
4 to 7	6
8 to 16	11
17 to 29	20
30 to 42	33
43 to 60	50
61 to 76	67
77 to 86	80
87 to 100	89

notice

 Now, multicast queue just support eight 1~67 effective ratio value on all platforms

Application notice



As show above: send packets from port1 and port2 to port3 in 10G speed in queue 0, and port3 will be in congestion.

For the multicast flow, if flow control is enabled on port1 and port2 and configure static threshold, the packets also perhaps will be dropped in port 3 egress queue. Ingress don't send pause frame because egress's shared-ratio is smaller than ingress's and egress reach it's threhold before igress reach.

So if users meet this issue, in this case, we suggest users configure dynamic mode. port1 and port2 will send out pause frame before port3 queue limit is reached, and accept burst packets to avoid packet drop.