MAD Congestion Control

The SA Management Datagrams (MAD) are General Management Packets (GMP) used to communicate with the SA entity within the InfiniBand subnet. SA is normally part of the subnet manager, and it is contained within a single active instance. Therefore, congestion on the SA communication level may occur.

Congestion control is done by allowing max_outstanding MADs only, where outstanding MAD means that is has no response yet. It also holds a FIFO queue that holds the SA MADs that their sending is delayed due to max_outstanding overflow.

The length of the queue is queue_size and meant to limit the FIFO growth beyond the machine memory capabilities. When the FIFO is full, SA MADs will be dropped, and the drops counter will increment accordingly.

When time expires (time_sa_mad) for a MAD in the queue, it will be removed from the queue and the user will be notified of the item expiration.

This features is implemented per CA port.

The SA MAD congestion control values are configurable using the following sysfs entries:

```
/sys/class/infiniband/mlx5_0/mad_sa_cc/
    1
        drops
        max_outstanding
        queue_size
        time_sa_mad
    2
    drops
    max_outstanding
    queue_size
    time_sa_mad
```

➤ *To print the current value:*

```
cat /sys/class/infiniband/mlx5_0/mad_sa_cc/1/max_outstanding 16
```

To

➤ *change the current value:*

```
echo 32 > /sys/class/infiniband/mlx5_0/mad_sa_cc/1/max_outstanding
cat /sys/class/infiniband/mlx5_0/mad_sa_cc/1/max_outstanding
32
```

To

➤ *reset the drops counter:*

```
echo 0 > /sys/class/infiniband/mlx5_0/mad_sa_cc/1/drops
```

Parameters' Valid Ranges

Parameter	Range		Default Values
	MIN	MAX	
max_oustanding	1	2^20	16
queue_size	16	2^20	16
time_sa_mad	1 milliseconds	10000	20 milliseconds