Rocky Enterprise Linux 9.2 Manual Pages on command 'tc-sfb.8'

$ man tc-sfb.8

SFB(8) Linux SFB(8)

NAME

sfb - Stochastic Fair Blue

SYNOPSIS

tc qdisc ... blue rehash milliseconds db milliseconds limit packets max packets target packets increment float decrement float penalty_rate packets per second penalty_burst packets

DESCRIPTION

Stochastic Fair Blue is a classless qdisc to manage congestion based on packet loss and link utilization history while trying to prevent non-responsive flows (i.e. flows that do not react to congestion marking or dropped packets) from impacting performance of responsive flows. Un? like RED, where the marking probability has to be configured, BLUE tries to determine the ideal marking probability automatically.

ALGORITHM

The BLUE algorithm maintains a probability which is used to mark or drop packets that are to be queued. If the queue overflows, the mark/drop probability is increased. If the queue becomes empty, the

probability is decreased. The Stochastic Fair Blue (SFB) algorithm is designed to protect TCP flows against non-responsive flows.

This SFB implementation maintains 8 levels of 16 bins each for account? ing. Each flow is mapped into a bin of each level using a per-level hash value.

Every bin maintains a marking probability, which gets increased or de? creased based on bin occupancy. If the number of packets exceeds the size of that bin, the marking probability is increased. If the number drops to zero, it is decreased.

The marking probability is based on the minimum value of all bins a flow is mapped into, thus, when a flow does not respond to marking or gradual packet drops, the marking probability quickly reaches one.

In this case, the flow is rate-limited to penalty_rate packets per sec? ond.

LIMITATIONS

Due to SFBs nature, it is possible for responsive flows to share all of its bins with a non-responsive flow, causing the responsive flow to be misidentified as being non-responsive.

The probability of a responsive flow to be misidentified is dependent on the number of non-responsive flows, M. It is (1 - (1 - (1 / 16.0)) ** M) **8, so for example with 10 non-responsive flows approximately 0.2% of responsive flows will be misidentified.

To mitigate this, SFB performs performs periodic re-hashing to avoid misclassification for prolonged periods of time.

The default hashing method will use source and destination ip addresses and port numbers if possible, and also supports tunneling protocols.

Alternatively, an external classifier can be configured, too.

PARAMETERS

rehash Time interval in milliseconds when queue perturbation occurs to avoid erroneously detecting unrelated, responsive flows as being part of a non-responsive flow for prolonged periods of time. Defaults to 10 minutes.

db Double buffering warmup wait time, in milliseconds. To avoid

destroying the probability history when rehashing is performed, this implementation maintains a second set of levels/bins as de? scribed in section 4.4 of the SFB reference. While one set is used to manage the queue, a second set is warmed up: Whenever a flow is then determined to be non-responsive, the marking proba? bilities in the second set are updated. When the rehashing hap? pens, these bins will be used to manage the queue and all non-responsive flows can be rate-limited immediately. This value determines how much time has to pass before the 2nd set will start to be warmed up. Defaults to one minute, should be lower than rehash.

limit Hard limit on the real (not average) total queue size in pack? ets. Further packets are dropped. Defaults to the transmit queue length of the device the qdisc is attached to.

max Maximum length of a buckets queue, in packets, before packets start being dropped. Should be slightly larger than target , but should not be set to values exceeding 1.5 times that of target . Defaults to 25.

target The desired average bin length. If the bin queue length reaches this value, the marking probability is increased by increment. The default value depends on the max setting, with max set to 25 target will default to 20.

increment

A value used to increase the marking probability when the queue appears to be over-used. Must be between 0 and 1.0. Defaults to 0.00050.

decrement

Value used to decrease the marking probability when the queue is found to be empty. Must be between 0 and 1.0. Defaults to 0.00005.

penalty_rate

The maximum number of packets belonging to flows identified as being non-responsive that can be enqueued per second. Once this

number has been reached, further packets of such non-responsive flows are dropped. Set this to a reasonable fraction of your uplink throughput; the default value of 10 packets is probably too small.

penalty_burst

The number of packets a flow is permitted to exceed the penalty rate before packets start being dropped. Defaults to 20 pack? ets.

STATISTICS

This qdisc exposes additional statistics via 'tc -s qdisc' output.

These are:

earlydrop

The number of packets dropped before a per-flow queue was full.

ratedrop

The number of packets dropped because of rate-limiting. If this value is high, there are many non-reactive flows being sent through sfb. In such cases, it might be better to embed sfb within a classful qdisc to better control such flows using a different, shaping qdisc.

bucketdrop

The number of packets dropped because a per-flow queue was full. High bucketdrop may point to a high number of aggressive, short-lived flows.

queuedrop

The number of packets dropped due to reaching limit. This should normally be 0.

marked The number of packets marked with ECN.

maxqlen

The length of the current longest per-flow (virtual) queue.

maxprob

The maximum per-flow drop probability. 1 means that some flows have been detected as non-reactive.

NOTES

SFB automatically enables use of Explicit Congestion Notification (ECN). Also, this SFB implementation does not queue packets itself. Rather, packets are enqueued to the inner qdisc (defaults to pfifo). Because sfb maintains virtual queue states, the inner qdisc must not drop a packet previously queued. Furthermore, if a buckets queue has a very high marking rate, this implementation will start dropping packets instead of marking them, as such a situation points to either bad con? gestion, or an unresponsive flow.

EXAMPLE & USAGE

To attach to interface $DEV, using default options:

tc qdisc add dev $DEV handle 1: root sfb

Only use destination ip addresses for assigning packets to bins, per? turbing hash results every 10 minutes:

tc filter add dev $DEV parent 1: handle 1 flow hash keys dst perturb 600

SEE ALSO

tc(8), tc-red(8), tc-sfq(8)

SOURCES

- W. Feng, D. Kandlur, D. Saha, K. Shin, BLUE: A New Class of Ac? tive Queue Management Algorithms, U. Michigan CSE-TR-387-99, April 1999.

AUTHORS

This SFB implementation was contributed by Juliusz Chroboczek and Eric Dumazet.