# CFF ID

## Proposed ID for use in CnTag

## August 2008 – Rev 1

**Caitlin Bestler**

**Caitlin.bestler@neterion.com**

# “Reaction Point ID” is not the ideal term

- **Hard to avoid implication that there is a Rate Limiter already in place.**
- **It really the Congestion Controllable Flow that we want to identify. The Rate Limiter comes later.**
  - Note: Congestion Control***lable*** Flow.
- **We *can* say that it reflects an inactive Rate Limiter**
  - But that only leads to confusion.
  - For many designs, all flows with the same Priority will be in a single queue *until* they are rate limited.

# End Station Output Queues

- **End Station Output Queues reflect many different design approaches:**
  - L2-only service, Offload/L4-L5 service, VM/Zone/Application specific, TCP vs UDP, …
    - And mixtures thereof
  - Multiple physical and/or virtual ports
  - Where memory lives: on-chip, on-host, external, etc.
  - What is in the queue:
    - TxDs versus Frames, mixtures (LSO).
    - Order of processing does not necessarily reflect theory.
- **DCB protocols must consider a large range of potential end station designs.**

# First Issue:
# Congestion Notification Message Scope

- **When an end station gets a CNM, which L2 flows should be rate limited?**
- **The CNM is already limited in scope**
  - Generated based on sampling at CP.
  - Unicast delivery back to a single end station.
- **But the CNM supplies information**
  - It is not a “speeding ticket”
  - Ideally all flows from this end station that reach the congested CP should be throttled
    - But what is realistic?
    - What set of frames should be impacted?

# Prior queuing should be Irrelevant

- **End stations have many designs**
  - Specific internal queue structures should neither be rewarded or penalized.
- **Frequently the pre-CNM queue will be too wide**
  - The end station will have had no reason to separate flows based on this destination.
  - Therefore many innocent flows will be slowed.
- **Sometimes the pre-CNM queue will be too narrow**
  - TOE/RDMA per-connection flows that are not the entire output from the end station to the destination.
- **Rate limited queues may be created *after* the CNM is received, the pre-CNM queue may fix relevant and irrelevant flows.**

# Therefore CFFs are not Queues

- **Rate Limited Flows MUST conform to the Rate Limiter.**
  - And they SHOULD cause minimal or no head-of-line blocking of other Flows.
- **A queue is certainly one method of achieving that**
  - But implementations must balance between benefits of multiple queues and their costs.
  - Placing Flows with similar Rate Limiters in the same Queue must be a valid option for implementations.
    - Especially for Flows without an active Rate Limiter.

# Proposed Definition of CFF ID

- **Each Flow is a member of a Flow Set**
  - There are only a small (TBD) number of Flow per Flow Set.
  - Multiple Flows per set are intended to support multi-pathing
    - They are not intended to reflect End Station internals.
    - Additional “source queue cookie” for internal use can be discussed separately.
- **Flow Set is determined by**
  - Egress Port
  - Destination: (VID + DA)
  - Priority
- **Additional L2-L4 Headers may be hashed to pick Flow within the Flow Set.**
  - There should be no effort expended to preserve order of frames that have different Flow IDs.
- **An End Station specific salt is then added to randomize the CFF ID.**
  - CFF_ID = f(Flow Set,multi-pathing-hash,end-station-salt)

# Changing Flow IDs

- **The prior definition could be extended to allow End Stations to include an administrative override and/or additional salt.**
  - This could be used by an end station to load balance its actual flows based on actual traffic patterns.
- **But, the Flow ID of any L4 flow MUST NOT be changed while the Flow ID is subject to a Rate Limiter.**

# Mouse/Elephant Problem

- **Nothing obligates the End Station to apportion the Rate Limiter evenly across all L4 Flows assigned the same Flow ID.**
  - The only requirement is that the Rate Limiter, as a whole, is complied with.
  - How the End Station allocates resources within the flows covered by a Rate Limiter is an implementation detail.