Flow control algorithms revisited

- Based on the Alladin, DVJ and Gandalf proposals

Stein Gjessing
Fredrik Davik

University of Oslo
NORWAY

Overview

- Based on the Alladin, DVJ, and Gandalf proposals:
  - Default and Excess Bandwidth Management
  - Max and min Passthru buffer configuration
  - Combination of Proactive and Reactive signaling

- Comparisons are based on **my understanding** of the Alladin, DVJ, and Gandalf proposals.
  Please correct me if I say something incorrect regarding these proposals

- Addendum: What does it mean to send at a specific rate ?

Default Bandwidth Allocation
Excess Bandwidth Usage

True fra all three proposals:

- All stations send at a Default/PreAllocated/ Provisioned Rate

- This is the same as if every station have a (pre) negotiated Weight

- **How to use/allocate excess bandwidth ?**

- **How to throttle/stop the use of excess bandwidth ?**

Default Bandwidth

True for all three proposals:

- The sum of all default rates does not congest the system (if configured correctly)

- Could be 0 for all

- Could be 1/N for all
  – where N is number of stations on the ring

- Could / should be partitioned into segments (by all stations – per VOQ)

VOQ: Virtual Output Queue

Passthru buffer configuration

Full station Passthru configuration
(Gandalf)

Pass A

Threshold 2   Threshold 1

Pass BC

| Priority and insert | High Buffer | Basic buffer |

Two Passthru fifos

A has (almost) absolute priority over BC

Passthru configuration
Minimum: One small Passthru fifo only
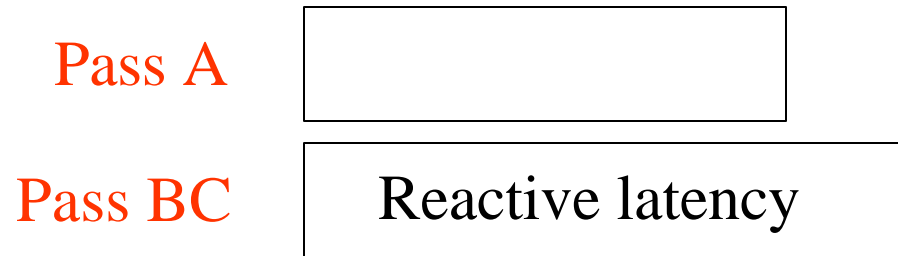### (Alladin)

Pass ABC

Insert

The buffer part of the passBC fifo has collapsed to nothing

The insert/prio of passBC and passA has collapsed into one

Lemma: All Passthru traffic has priority over Add traffic

Passthru configuration

In between example: Two Passthru fifos
(DVJ)

Pass A

Pass BC   Reactive latency

Conclusion on Passthru config.

- The Alladin, DVJ and Gandalf proposals can be viewed as three versions of a general Passthru buffer configuration specification

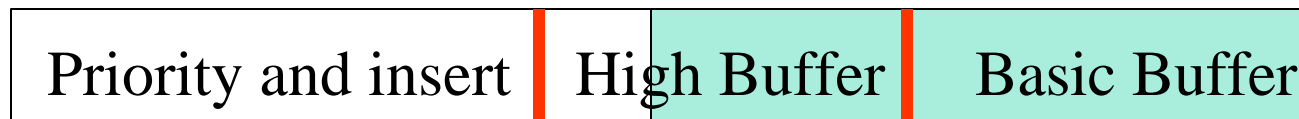# Flow control/ Fairness algorithm

(What we want)

- The same algorithm for both max and min Passthru configurations (and all in-between configurations)

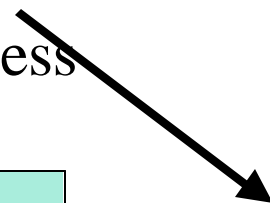- The same set of control messages for all configurations

Station add traffic

- As long as there are packets in the Buffer part of the passthru fifo, and not in the prio&insert part, the station adds traffic with its default rate
  and may send excess traffic

Add/Transmit
@ default rate
+ excess

Pass BC | Priority and insert | High Buffer | Basic Buffer

Station add traffic

(max configuration)

- As long as there are no packets in the passthru BC fifo, the station adds BC traffic with its default rate
  and may send excess traffic

Add/Transmit
@ default rate
+excess

Pass BC

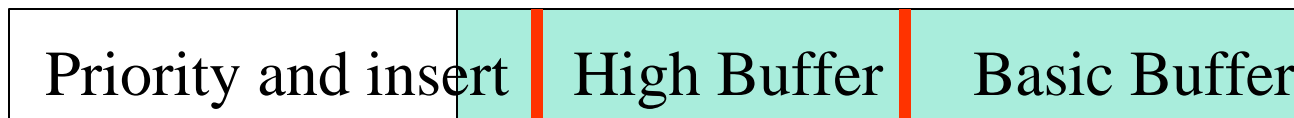| Priority and insert	High Buffer	Basic Buffer

Station add traffic
(max configuration)

- As long as there are packets in the Priority and insert part of the passthru fifo, the station does not add traffic (except for A traffic)

No BC added,
A only

Pass BC | Priority and insert | High Buffer | Basic Buffer →

Station add traffic
(min configuration)

- As long as there are no packets in the passthru fifo, the station adds traffic with its default rate
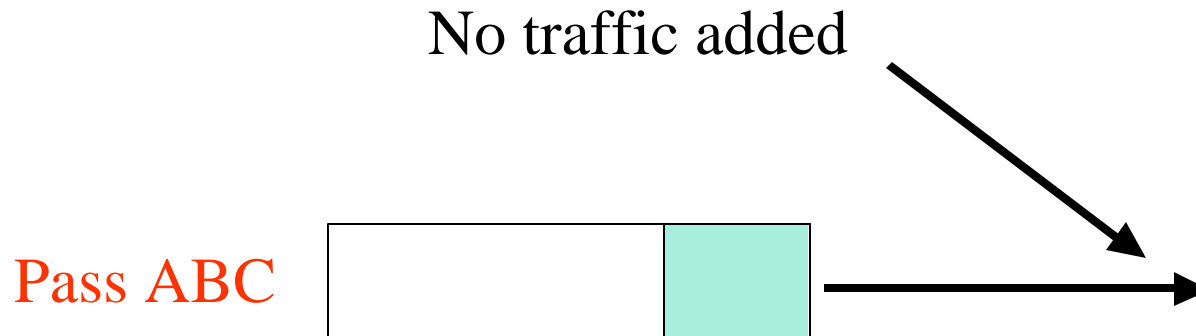 *and may send excess traffic*

Add/Transmit
@ default rate
+excess

Pass ABC

802-17-01-00111, sg_fctrl_03.pdf Stein Gjessing, Fredrik Davik

Station add traffic
(min configuration)

- As long as there are packets in the passthru fifo, the station does not add any traffic

No traffic added

Pass ABC

The ultimate questions

- When can a station send excess traffic?
- How to allocate and de-allocate excess traffic

- By   Clients ?
- By   MAC ?

- **Proactive** Congestion Avoidance (by Client ?) and/or
- **Reactive** Congestion Detection and Relief (by MAC ?)

Proactive = Congestion avoidance

- Notify other stations about available excess bandwidth

- Notify again when this bandwidth is not available any more

- Notification decided by Clients

- Enforced by MACs

(Alladin)

Reactive = Congestion Detection and Relief

- Congestion detection
  - PassBC is filled above Threshold
- Relief
  - Send usage rate or idle count upstream

Gandalf   **DVJ**

- Decided and enforced by MAC

Congestion management

- Proactive is a more long term scheme (Clients)
- Reactive is a more short-term / dynamic scheme (MAC)
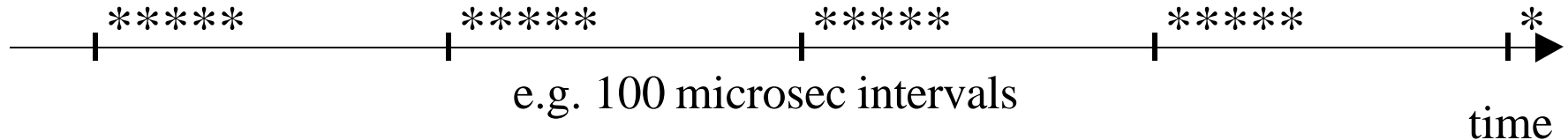
Stations should be
- Proactive if Client have the information (Alladin)
- Reactive if information is not available (Gandalf and DVJ)

- Can we have both ???
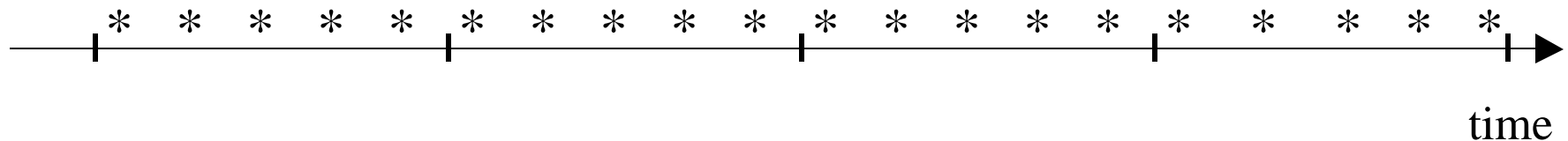 without creating a camel or a dinosaur

Conclusion

- Both proactive and reactive excess bandwidth management is needed
- Must be seen in connection with station passthru configuration

- I believe proactive and reactive management can and should interoperate

- Control messages for both needed
- Clarification of interoperability needed

Addendum: What does it mean to send at a specific rate

- The stations should send packets ALL the time according to its allowed rate
- Means: Not in bursts
- E.g. not in 100 microsecond bursts like:

```
     *****        *****        *****        *****        *
  ───┼────────────┼────────────┼────────────┼────────────►
           e.g. 100 microsec intervals
                                                       time
```

- But like:

```
   *  *  *  *  * *  *  *  *  * *  *  *  *  * *  *  *  *  *
  ─┼────────────┼────────────┼────────────┼────────────►
                                                      time
```

* Means one packet sent