Urgency Based Scheduler
- updates since March 2013 -

Johannes Specht (Univ. Duisburg-Essen)
Background and Content

• **Content**  Urgency Based Scheduler (UBS)
  (based on March 2k13 slides, changes based on feedback)

• **UBS**  Flexible Traffic Class?
  (cmp. AAA2C slides)

**UBS Concepts:**

• Sub Shapers and Sub Queues
• Priority Queues
• Latency Calculation
• Ingress Policing
• Higher and Lower Traffic Classes
• Scheduling

16.10.2013  Johannes Specht - University of Duisburg-Essen
AVB1 – THIS MAKE’S LATENCY CALCULATION TRICKY
Example AVB1 (1)

**Example scenario - simplified environment**
- Fast Ethernet, two bridges, store & forward
- One AVB class (CBSA), no other classes → no need for traffic selection, etc.
- No further delays (link delays, etc.)
- **Red traffic**: One (or more!) streams, one path, constant frame size

**Focus: Shaper to shaper subsystem (1 Hop)**
- **Shaper S1**: Guaranteed input bound
- **Shaper S2**: Guaranteed output bound

**Max. Latency: Isolated subsystem analysis possible!**
- input bound ≈ output bound:
  → “Entering traffic can leave”...
Example AVB1 (2)

Extended Scenario: as before, but...
- Green traffic: One (or more!) streams, one path, equal frame size
  Note: Reservations are not forced to send!

Max. Latency:
Isolated subsystem analysis impossible!

- Input bound > output bound
  \(\Rightarrow\) “Entering traffic can’t leave”...
- Depends on history of the traffic (topology, other traffic)
  \(\Rightarrow\) Complicated \(\Rightarrow\) Undesired
SUB SHAPERS AND SUB QUEUES
Sub Shapers – Purpose

**Always assure:**

\[ input \text{ rate} = output \text{ rate} \]

**Enables:**

- Maximum latency calculation per subsystem

**Sub Queues:**

- Consequence of Sub Shapers (1 Sub Queue per Sub Shaper)

Notes for correctness:
- (1) Input rate < output rate is also ok
- (2) CBSA, but without being blocked by other frames/classes

16.10.2013

Johannes Specht - University of Duisburg-Essen
**Sub Shapers – Cont.**

*Interfering Sub Shapers:*
- Guaranteed input bound
- Maximum latency calculation per subsystem possible
Sub Queues: (regular) FIFO Queues

Class Queue: Priority Queue

PrioriTBY Queue

UBS (egress)
**Problem:** Latency → Function of Hops

**Simplified**
- More hops ≈ more End-2-End latency
- **but**
- Maximum latency *is a requirement* ...
- ... *not only* a function of the topology

**Example:**

<table>
<thead>
<tr>
<th>Traffic</th>
<th># Hops</th>
<th>Rate</th>
<th>Max. Required End-2-End Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>7</td>
<td>25 Mbit/s</td>
<td>1 ms → Urgent</td>
</tr>
<tr>
<td>Green</td>
<td>2 or 3</td>
<td>25 Mbit/s</td>
<td>10 ms → Relaxed</td>
</tr>
</tbody>
</table>
Priority Queue – Purpose

Accelerate Urgent Traffic

Example:

<table>
<thead>
<tr>
<th>Traffic</th>
<th># Hops</th>
<th>Rate</th>
<th>Max. Required End-2-End Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>7</td>
<td>25 Mbit/s</td>
<td>1 ms → Urgent</td>
</tr>
<tr>
<td>Green</td>
<td>2 or 3</td>
<td>25 Mbit/s</td>
<td>10 ms → Relaxed</td>
</tr>
</tbody>
</table>

Queue urgent traffic “before” more relaxed traffic, i.e. “make red traffic faster”
**Priority Queue - Example**

**Accelerate Urgent Traffic**

<table>
<thead>
<tr>
<th>Traffic</th>
<th># Hops</th>
<th>Rate</th>
<th>Max. Required End-to-End Latency</th>
<th>Max. Frame Size</th>
<th>Max. Calculated End-to-End Latency*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>7</td>
<td>25 Mbit/s</td>
<td>1 ms → urgent</td>
<td>128 Byte</td>
<td>962,56 µs @ 7 Hops</td>
</tr>
<tr>
<td>Green</td>
<td>3 or 2</td>
<td>25 Mbit/s</td>
<td>10 ms → relaxed</td>
<td>512 Byte</td>
<td>1331,20 µs @ 3 Hops</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>832,85 µs @ 2 Hops</td>
</tr>
</tbody>
</table>

**Assumptions**
- 100 Mbit/s Fast Eth.
- No TDMA Class present
- No link delays
- Preemption of lower classes (max. 128 Byte framelet)
- Simple priority queue algorithm: Strict priority per Sub-Shaper/-Queue

*Result of a preliminary worst case calculation

16.10.2013 Johannes Specht - University of Duisburg-Essen
Summary

Proposal for Flexible Traffic Class

- Urgency Based Scheduling
  - Sub shaper & Sub queues:
    - Enables latency analysis per Hop
  - Priority queue:
    - Loosens topology dependency on max. Latency
    - Maps Latency Requirements

Next steps
- Present Latency Calculation
- Other?
Thank you for your Attention!

Questions, Opinions, Ideas?

Johannes Specht
Dipl.-Inform. (FH)

Dependability of Computing Systems Institute for Computer Science and Business Information Systems (ICB) Faculty of Economics and Business Administration University of Duisburg-Essen

Schützenbahn 70 Room SH 502
45127 Essen GERMANY

T +49 (0)201 183-3914
F +49 (0)201 183-4573

specht@dc.uni-due.de
http://dc.uni-due.de