Efficient Transport of Isochronous Streams in Residential Ethernet:
\textit{With A Generalized Admission Control Approach}

Felix Feifei Feng
feng.fei@samsung.com

Sihai Wang
sihai.wang@samsung.com

Samsung Electronics
Background
Terminology

- **Subscription**
  - End-to-end procedure of setting up an isochronous stream. It will employ the Simple Reservation Protocol (SRP) which is the signaling for conveying resource reservation information.

- **Admission control**
  - Local operation in each bridge, which assesses whether there are enough resources locally in this bridge to support corresponding triggering reservation signaling.

- **Pacing**
  - By holding each isochronous frame until its corresponding issuing time, pacing mechanism maintains the traffic pattern along the stream path to guarantee a low jitter bound and evenly distributed buffer space requirement inside the network. More specifically with current ResE approach [1,2]:
    - “The classA frames are gated to prevent their early departure. Gating involves blocking classA frames that arrived with sourceCycle=n, until the start of cycle n+p. After the start of cycle n+p, the transmitter waits for the completion of preceding non-classA frames (or residual cycle n+p-1 classA frames), then transmits these arrived-in-cycle-n frames with sourceCycle=n+p.”
    - Delay, jitter and buffer requirement in each bridge is bound to pacing parameter p
Traffic Distortion

- A switch can provide local performance guarantees to a stream only when its traffic pattern satisfies certain specifications.
- But interactions between different isochronous streams and asynchronous frames may distort the traffic pattern of a stream.
- The distortion may cause violation of the specification even if the stream satisfies the specification at the entrance to the network.
Rate-controlled Static Priority Queuing [1]

- A rate controlled server has two components: a rate controller and a scheduler
  - Rate controller shapes input stream traffic to desired traffic pattern by assigning an eligibility firing time to each frame
  - Scheduler orders the transmission of eligible frames

Figure 1: Rate-Controlled Static-Priority Queueing
Application of RCSP to ResE

- We can use the delay-jitter controlling regulator model
  - The scheduler define the maximum possible delay experienced by a isochronous frame
    - Delay bound is derived from traffic profile and admission control criteria
  - For each incoming isochronous frame, the regulator will compensate the difference between the maximum possible delay and the frame’s actually experienced delay in the upstream node
    - Jitter will not be accumulated.
    - Maximum buffer requirement can be fixed.

- We defined a base cycle for all isochronous streams in ResE, then the rate-controller (regulator) can be simplified:
  - Relative eligibility time can be used instead of absolute eligibility time
    - “The classA frames are gated to prevent their early departure. Gating involves blocking classA frames that arrived with sourceCycle=n, until the start of cycle n+p. After the start of cycle n+p, the transmitter waits for the completion of preceding non-classA frames (or residual cycle n+p-1 classA frames), then transmits these arrived-in-cycle-n frames with sourceCycle=n+p.”
  - Since all streams use a same cycle, there is no need for per-connection regulation
Efficient Transport of Isochronous Streams
Problem Statement: Low Bandwidth Streams

- In 100M Ethernet a prevalence choice of cycle size is 125us, and the pacing parameter $p$ is set to 2. [2,3]
  - Since all streams reserve identical bandwidth in each 125us cycle, for low bandwidth streams it causes the problem of high overhead.
    - The overhead of transmit a payload data unit in Ethernet includes IPG (12bytes), Preamble(8bytes), DA(6bytes), SA(6bytes), Length/Type(2bytes), FCS(4bytes), Pad(0~46bytes, depend on PDU size)
    - For example, for a 2Mbps CD audio stream
      - With a 125us cycle, The utilization is ~38% (32bytes payload per cycle, $32/(32+38+14)\approx38\%$)
    - But by using a large cycle size, the utilization can be obviously improved
      - With a 250us cycle, the utilization is ~62% (63bytes payload per cycle, $63/(63+38)$)
      - With a 500us cycle, the utilization is ~77% (125bytes payload per cycle, $125/(125+38)$)
  - We need a solution to improve the utilization efficiency for low bandwidth streams while keeping same low delay/jitter performance and implementation simplicity as previous approach.
Existing Solution

- Encapsulate several low bandwidth content into blocks [2]
  - However, the application of this scheme is limited:
    - Only content streams with same source-destination nodes pair can be multiplexed
    - Source node and destination node need additional processing power for encapsulation and de-encapsulation.

![Diagram showing encapsulation and de-encapsulation of blocks within a frame]

c) Groups of blocks
Proposed Solution

- **Sub-rate allocation**
  - This concept was defined in IEEE802.15.3 as “A channel time allocation that occurs only once every n superframes (n>1)”.

- **Sub-rate traffic model in ResE**
  - A base-cycle of $P_{\text{base}}$ is defined for ResE system
  - Isochronous streams can use a value which is the multiple of base-cycle ($P_{\text{base}}$) as their stream data traffic period.
    - Those streams send out frames periodically based on their individual cycle.
      - Low bandwidth streams can then fit into the sub-rate traffic model to keep its utilization ratio above a reasonable threshold.
    - For the implementation feasibility, the maximum number of cycle classes in ResE system should be limited. For example, 2~4 classes would be effective and operational.

![Diagram showing mono-cycle traffic and multi-cycle traffic](image-url)
Proposed Solution (cont.)

For the sub-rate traffic model scenarios, a more generalized admission control criteria should be used to guarantee the same delay/jitter performance for isochronous stream packets as fixed rate scenarios in previous approach while using the same isochronous packets pacing scheme.

- Notations:
  - \( N \), the maximum allowable number of traffic classes in system
  - \( B_n \) (\( n=1...N \)), the maximum cumulative bandwidth (bit/s) that is allowed to be assign to class-\( n \) traffic. The bandwidth should include all possible overhead.
  - \( C_n \) (\( n=1...N \)), the ratio of class-\( n \) traffic’s period to the basic-cycle.
  - \( r \), the ratio of total link capacity that can be assigned to isochronous streams.
  - \( l \), link capacity
  - \( k \), the transmit time (include all overhead) for a largest asynchronous packet. (For example, in 100M Ethernet, \( k \sim 123 \text{us} \); in GbE, \( k \sim 12.3 \text{us} \))
  - \( p \), pacing parameter. Isochronous frames that arrived cycle \( m \) will be paced to be forwarded utile cycle \( m+p \).
Proposed Solution (cont.)

The generalized admission control condition is as follows [1]:

- Maximum accumulative allocated bandwidth constraint:
  \[
  \sum_{n=1}^{N} B_n \leq r \times l
  \]  
  (1)

- Worst case local delay should be less than the pacing holding time (\(F_{recovery}\), \(w_{asme}\), \(h_{ilk}\) can be ignored. This delay can be easily taken into account by adding its worst-case value to the right side of equation(2))

\[
\sum_{n=1}^{N} \frac{B_n \times C_n \times P_{base}}{l} + k \leq p \times P_{base}
\]  
  (2)

or be rewritten as:

\[
\sum_{n=1}^{N} C_n B_n \leq (p - \frac{k}{P_{base}}) \times l
\]  
  (3)

By substituting system parameters into above equations, admission control parameters \(B_n\) can be calculated.
The admission control condition can also be easily written in other forms:

- In the form of cumulative allocated time-slice ($T_n$, including all overhead) per sub-rate cycle:

$$\sum_{n=1}^{N} \frac{T_n}{C_n} \leq r \times P_{base}$$

$$\sum_{n=1}^{N} T_n \leq p \times P_{base} - k$$

- Similarly, it can be written in the form of cumulative allocated bits ($b_n$, including all overhead) per sub-rate cycle:

$$\sum_{n=1}^{N} \frac{b_n}{C_n} \leq r \times l \times P_{base}$$

$$\sum_{n=1}^{N} b_n \leq (p \times P_{base} - k) \times l$$
Extension of information carried in subscription signaling

- To support sub-rate traffic model, a new field Traffic_Rate_Factor should be added to existing subscription information
  - Traffic_Rate_Factor is an integer that equals to the ratio of the traffic frames’ period to the base-cycle, which is notated as Cn in the admission control conditions.
  - Totally two input parameters are used in admission control: sub-rate class and bandwidth requirement (Cn and Bn)
Proposed Solution (cont.)

- **System operation**
  - Based on the application requirements, system parameters N, Cn, r and p are predefined for the ResE network. Then corresponding admission control parameters Bn are determined and configured.
    - Bn may even be adaptively adjusted in a ResE network
  - When an application needs to set up a new class-n isochronous stream, it indicates its traffic class n and its bandwidth requirement Bq in the subscription protocol signaling.
  - Each relevant ResE switch makes admission control decisions by comparing if the accumulative bandwidth for class-n traffic will exceed the corresponding admission control parameter Bn.
    - If not, the admission control is successful on this switch
    - Otherwise, the admission control is failed on this switch
Design Examples: Fast Ethernet

Assumptions:
- 100Mbps Ethernet
- Two kinds of CBR streams. Base cycle $P_{\text{base}}$ is set to 125us. One of the stream uses 125us cycle; the other one uses 250us cycle.
- At most 75% link capacity can be allocated to those streams.
- Pacing parameter $p$ is set to 2.
- Note the maximum bandwidth can be allocated to the two kinds of streams as $B_{125}$ and $B_{250}$, respectively.

Then the admission control conditions are:

$$\begin{cases} 
B_{125} + B_{250} \leq 75\text{Mbps} \\
B_{125} + 2 \times B_{250} \leq 100\text{Mbps}
\end{cases}$$

By using a 250us cycle, a 2Mbps CD audio can increase its utilization ratio from 38% to 62% compared with conventional 125us cycle model.
Design Examples: GB Ethernet

Assumptions:
- 1000Mbps Ethernet
- Two kinds of CBR streams. Base cycle \( P_{\text{base}} \) is set to 125us. One of the stream uses 125us cycle; the other one uses 250us cycle.
- At most 75% link capacity can be allocated to those streams.
- Pacing parameter \( p \) is set to 1.
- Note the maximum bandwidth can be allocated to the two kinds of streams as \( B_{125} \) and \( B_{250} \), respectively.

Then the admission control conditions are:

\[
\begin{align*}
B_{125} + B_{250} &\leq 750Mbps \\
B_{125} + 2 \times B_{250} &\leq 900Mbps
\end{align*}
\]

By using a 250us cycle, a 2Mbps CD audio can increase its utilization ratio from 38% to 62% compared with conventional 125us cycle model.
Further Extension: VBR Stream Scenario

Problem statement:

- Some applications in ResE will use VBR traffic model
- Current solution reserves resources only based on the peak rate
  - It may decrease the acceptance ratio of admission control
Proposed Solution

Each admission control request includes both its average bandwidth requirement and its peak bandwidth requirement.

- Cumulative allocated average bandwidth should be less than a fixed ratio to protect the performance of asynchronous data
- Cumulative allocated peak bandwidth should be less than the line rate and guarantee the delay bound
Proposed Solution (cont.)

We can derive the admission control conditions (in the form of allocated bits per sub-rate cycle) [1]:

\[
\begin{align*}
\sum_{n=1}^{N} \frac{ba_n}{C_n} & \leq r \times l \times P_{\text{base}} \\
\sum_{n=1}^{N} bp_n & \leq (p \times P_{\text{base}} - k) \times l \\
\sum_{n=1}^{N} \frac{bp_n}{C_n} & \leq l \times P_{\text{base}} \\
ba_n & \leq bp_n \quad (n = 1 \ldots N)
\end{align*}
\]

- \( ba_n \) (n=1…N): the maximum cumulative bits in terms of average that is allowed to be assigned to class-n traffic, including all possible overhead.
- \( bp_n \) (n=1…N): the maximum cumulative bits in terms of peak that is allowed to be assigned to class-n traffic, including all possible overhead.

→ Totally three input parameters are used in admission control: sub-rate class, average bandwidth requirement, and peak bandwidth requirement \((C_n, ba_n \text{ and } bp_n)\)
Design Examples

100Mbps; $P_{\text{base}} = 125\text{us}$; No sub-rate streams; $p=2$; 75% for Iso.

\[
\begin{align*}
ba_{125} & \leq 9375(\text{bits}) \\
bp_{125} & \leq 12696 \approx 12500(\text{bits}) \\
bp_{125} & \leq 12500(\text{bits}) \\
ba_{125} & \leq bp_{125}
\end{align*}
\]

100Mbps; $P_{\text{base}} = 125\text{us}$; Two kinds of streams; $p=2$; 75% for Iso.

\[
\begin{align*}
ba_{125} + \frac{ba_{250}}{2} & \leq 9375(\text{bits}) \\
bp_{125} + bp_{250} & \leq 12696 \approx 12500(\text{bits}) \\
bp_{125} + \frac{bp_{250}}{2} & \leq 12500(\text{bits}) \\
ba_{125} & \leq bp_{125}; ba_{250} \leq bp_{250}
\end{align*}
\]

\[
\begin{align*}
ba_{125} & = 9000(\text{bits}) \\
ba_{250} & = 750(\text{bits}) \\
bp_{125} & = 10000(\text{bits}) \\
bp_{250} & = 2500(\text{bits})
\end{align*}
\]
Design Examples (cont.)

- **1000Mbps;** $P_{\text{base}} = 125\text{us};$ No sub-rate streams; $p=2; 75\%$ for Iso.

  \[
  \begin{align*}
  ba_{125} &\leq 93750(\text{bits}) \\
  bp_{125} &\leq 237696(\text{bits}) \\
  bp_{125} &\leq 125000(\text{bits}) \\
  ba_{125} &\leq bp_{125}
  \end{align*}
  \]

- **1000Mbps;** $P_{\text{base}}=125\text{us};$ Two kinds of streams; $p=2; 75\%$ for Iso.

  \[
  \begin{align*}
  ba_{125} + \frac{ba_{250}}{2} &\leq 93750(\text{bits}) \\
  bp_{125} + bp_{250} &\leq 237696(\text{bits}) \\
  bp_{125} + \frac{bp_{250}}{2} &\leq 125000(\text{bits}) \\
  ba_{125} &\leq bp_{125} ; ba_{250} &\leq bp_{250}
  \end{align*}
  \]

\[
\begin{align*}
ba_{125} &= 75000(\text{bits}) \\
ba_{250} &= 37500(\text{bits}) \\
bp_{125} &= 100000(\text{bits}) \\
bp_{250} &= 50000(\text{bits})
\end{align*}
\]
Further Extension: Multi-Priority Iso. Streams

- In previous solutions, all isochronous streams belong to a same priority class, which means all streams share the same per-hop delay and jitter performance.
  - Per-hop Delay = p*P_{base}
  - End-to-end Jitter = p*P_{base}

- ResE applications may need differentiated performance between streams.
  - Isochronous streams may be classed with different priorities. Each priority class provides its specific per-hop delay and end-to-end jitter performance value.
Current Solutions

- **A rate-monotonic scheduling based method [2]**
  - "Rate-based scheduling involves associating a priority with frame transmissions, where the priority is a monotonic function of the frame transmission frequency"

- **Disadvantage:**
  - The performance of isochronous streams is bound with their transmission frequency. But real applications’ requirement can be more general. There is not necessarily relationship between the delay/jitter performance and the transmission frequency.
Proposed Solutions

Isochronous streams are assigned with priority classes.

- Assume there are Q priority classes. Subscript q (q = 1...Q) is used to indicate the priority classes.
- For each priority class q, a specific pacing parameter p^q is used.
  - Frames of priority class q with arrival-cycle n will be gated to prevent their early departure. They aren’t eligible for departure until the start of cycle n+p^q.
  - p^1<p^2<...<p^Q
  - Therefore specific per-hop delay (p^q*P_{base}) and end-to-end jitter (p^q*P_{base}) performance is provided for this class
- Non-preemptive strict priority queues are used for the forwarding of eligible frames
  - Class 1 has the highest priority; Class Q has the lowest priority

In each priority class, the isochronous streams can still use different transmission frequency (sub-rate streams) and VBR traffic model.
Proposed Solutions

Corresponding admission control conditions can be derived as [1]:

\[
\begin{align*}
\sum_{q=1}^{Q} \sum_{n=1}^{N^q} \frac{ba_n^q}{C_n^{q}} &\leq r \times l \times P_{\text{base}} \\
\sum_{q=1}^{m} \sum_{n=1}^{N^q} bp_n^q + p^m \times \sum_{q=1}^{m-1} \sum_{n=1}^{N^q} \frac{bp_n^q}{C_n^{q}} &\leq (p^m \times P_{\text{base}} - k) \times l \quad (m = 1...Q) \\
\sum_{q=1}^{Q} \sum_{n=1}^{N} \frac{bp_n^q}{C_n^{q}} &\leq l \times P_{\text{base}} \\
ba_n^q &\leq bp_n^q \quad (n = 1...N; \ q = 1...Q)
\end{align*}
\]

→ Totally four input parameters are used in admission control: priority class, sub-rate class, average bandwidth requirement and peak bandwidth requirement (p^q, C_n^{q}, ba_n^q and bp_n^q)
Conclusion

This extended traffic model and corresponding generalized admission control criteria provide us a ResE approach which can:

- use the same simple pacing mechanism as previous approach.
- keep the same low delay/jitter performance as previous approach
- improve the utilization efficiency for low bandwidth streams
- improve the acceptance ratio for VBR streams
- provide priority differentiation for isochronous streams. And there is no limitation on the relationship between the delay/jitter performance and transmission frequency.
Reference

[1] H. Zhang et. al in: *Rate-controlled static priority queuing*, Infocom93

[2] Residential Ethernet (RE) (a DVJ working paper),
   http://grouper.ieee.org/groups/802/3/re_study/material/index.html

[3] Residential Ethernet Tutorial,
   http://grouper.ieee.org/groups/802/3/tutorial/mar05/tutorial_1_0305.pdf
Thanks