A Proposal for RPR Rate Synchronization

July 2002, Vancouver, BC

Rhett Brikovskis, Lantern Communications
David James, Cypress Semiconductor
John Lemon, Luminous Networks
Mike Takefman, Cisco Systems

Overview

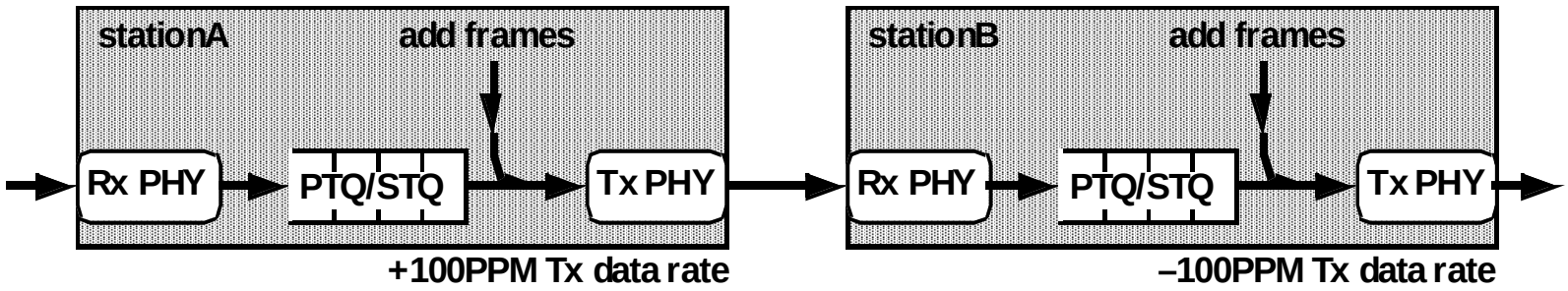
- What is rate synchronization and why is it needed?
- An example of the rate synchronization problem in RPR
- How can rate synchronization be performed?
- Proposal for a simple rate synchronization function
- Annex B issues
- Open issues
- Summary and recommendations

What Is Rate Synchronization

- RPR stations may use either synchronous or asynchronous PHYs.
- In a synchronous ring, the transmit data rate at each station is exactly identical to the received data rate. A station cannot receive more traffic than it is able to transmit.
- In an asynchronous ring, the transmit data rate at each station is determined by a local clock source, and varies slightly from the nominal link rate. If a station transmits at a lower data rate than the preceding station, the transit queue may fill.
- This proposal describes a rate synchronization mechanism to prevent PTQ overflow.
- The rate synchronization discussed here is not:
  - A clock synchronization function
  - Bit-synchronous operation

An Example

- A simplified example is shown below.
  - Station A is transmitting at the nominal link rate plus 100 PPM.
  - Station B is receiving only transit traffic from Station A, but is transmitting only at the nominal link rate minus 100 PPM.
  - If the link is fully utilized, PTQ in B will fill at (link rate)*200 PPM.
  - If this condition is sustained for a sufficient period of time, the PTQ in B will overflow and lose packets.



How To Resolve RX/TX Rate Differences?

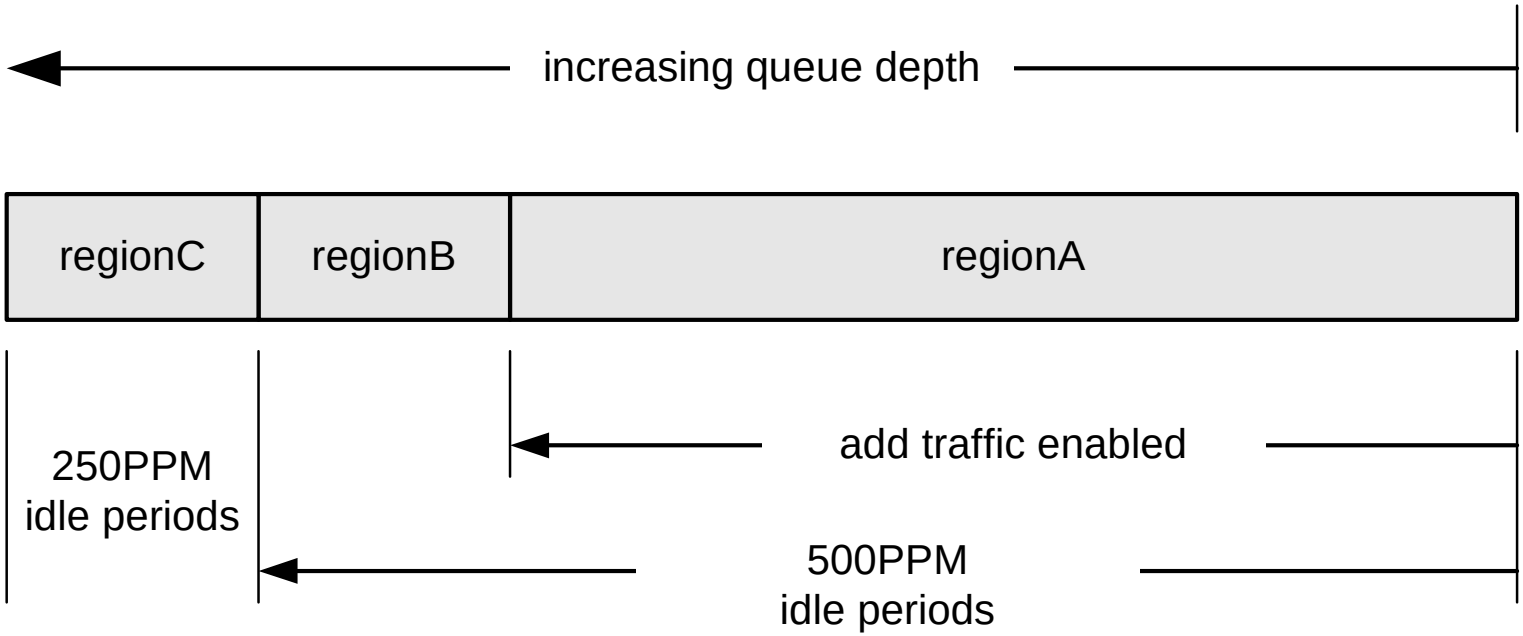
- Why not use RPR fairness to provision the link for <100%?
  - Fairness operates on a time-averaged basis, and has a non-zero response time. Fairness cannot ensure that link utilization won't burst to 100% for long enough to overflow the PTQ. Link also may be over-provisioned and reach full utilization. Provisioning guidelines may resolve this issue in specific applications.
- Why not force the link utilization to <100% by inserting small idle periods (similar to Ethernet IPG)?
  - When link reaches maximum utilization, idle periods will be inserted to limit data rate to <100%. However, the following station must still re-transmit the same minimum idle periods it receives, and the receive/transmit rate difference is unchanged. The PTQ will fill.
- Need to insert small, variable idle periods in the transmitted datastream, and vary their size based on the TX/RX rate difference.

A Simple Rate Synchronization Function

- Insert small idle periods into the transmitted datastream of each RPR station, and vary the time interval between the idle periods in accordance with the depth of the PTQ:
  - Any time the PTQ level is below a first threshold (PTQ occupancy is low), insert fixed idle periods at a regular interval corresponding to 500 PPM of the nominal link data rate.
  - Above the first threshold, inhibit the addition of any local traffic (add traffic or control packets). This ensures that the PTQ can only continue to fill if the RX rate exceeds the TX rate and link utilization is high.
  - Above a second PTQ threshold, increase the interval between idle periods to a rate corresponding to 250 PPM of the nominal link data rate. This increases the TX data rate to reduce the PTQ level.

A Graphical View

- Operation of the algorithm vs PTQ depth



Generating the Idle Period

- To generate idle periods, the MAC could transmit nothing (same as between frames in a link that is less than fully utilized), but these idle periods will not be preserved if the MAC and PHY are in different clock domains.
- Propose an “Idle Packet” instead that accommodates all RPR PHYs:
  - Define a “Type 0” Fairness Message variant called an Idle Packet.
  - Idle Packet is a point-to-point message only, and is discarded by recognition or by TTL.
  - Idle Packet is a fixed 16-byte length with a fixed payload value.

What is Annex B?

- Annex B is titled “Transmit Clock Synchronization”. No proposals have been made to synchronize clocks. As a minimum, this annex should be re-titled as “RPR station rate synchronization”.
- The function described in this proposal is part of the MAC datapath. The rate synchronization sub-ad-hoc believes it is better suited to inclusion in Clause 6 than in a separate annex. Annex B should be removed.

Open Issues

- The rate synchronization sub-ad-hoc agrees on the general concepts presented here, but has not closed several [important] issues:
  - Is this an optional or mandatory requirement? Should synchronous implementations (for example) be burdened with implementing this function even though unneeded.
  - Need a draft text proposal. First-cut text was just completed (NewRate.pdf) but is only now being reviewed by the sub-ad-hoc.
  - Are the 250 PPM and 500 PPM idle periods that have been selected appropriate under all conditions?

Summary and Recommendations

- A method is needed to ensure that the PTQ cannot overflow in asynchronous RPR implementations.
- A simple rate synchronization function is defined to insert small idle periods in the transmit datastream, and vary the rate based on PTQ level.
- Sub-ad-hoc proposes to refine a detailed text proposal for a rate synchronization function for WG review. Anyone with interest is encouraged to participate. Initial draft text is available as NewRate.pdf.
- Sub-ad-hoc expects to incorporate proposed text in Clause 6 and delete Annex B.