MAC Efficiency Issues

Marek Hajduczenia, PhD
ZTE Corporation
marek.hajduczenia@zte.pt

Motivation

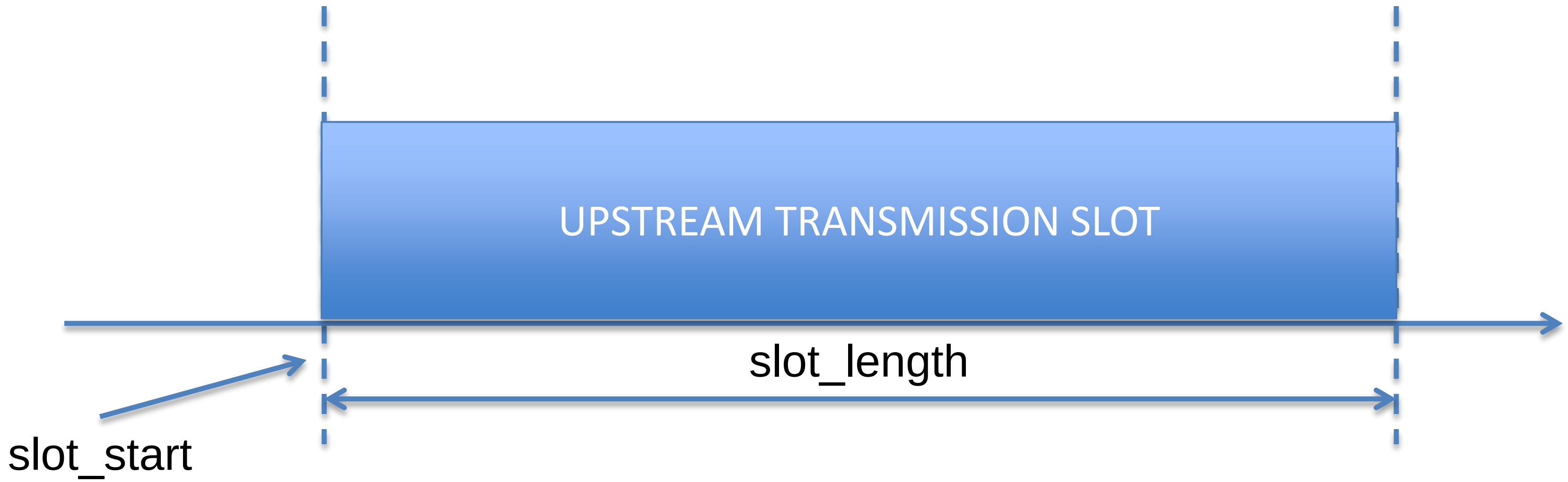
- A lot of energy has been spent so far discussing efficiency of upstream channel in EPoC
- The goal: design the system with maximum upstream efficiency to haul as many bits as possible
- However, so far we have limited ourselves to PHY layer efficiency problems
- This contribution looks at the problem of upstream scheduling and MAC-level efficiency, drawing conclusions on EPON studies available in available publications

Upstream transmission in EPON

- Upstream transmission in EPON is bursty and requires centralized scheduling
- Sometimes ONU does not fill in the allocated timeslot completely, resulting in bandwidth loss (underutilization).
- There are several sources of upstream inefficiencies in EPON, including:
  - For heavy load: slot remainder, imprecise reporting / bandwidth granting , changes in frame boundary alignment between the times of reporting and granting
  - For light load: changes in frame boundary alignment between the times of reporting and granting
  - TQ granularity: 1 TQ = 2 octets in 1G-EPON and 20 octets in 10G-EPON, affecting granting and reporting boundaries (always present)
- Individual items are discussed on the following slides

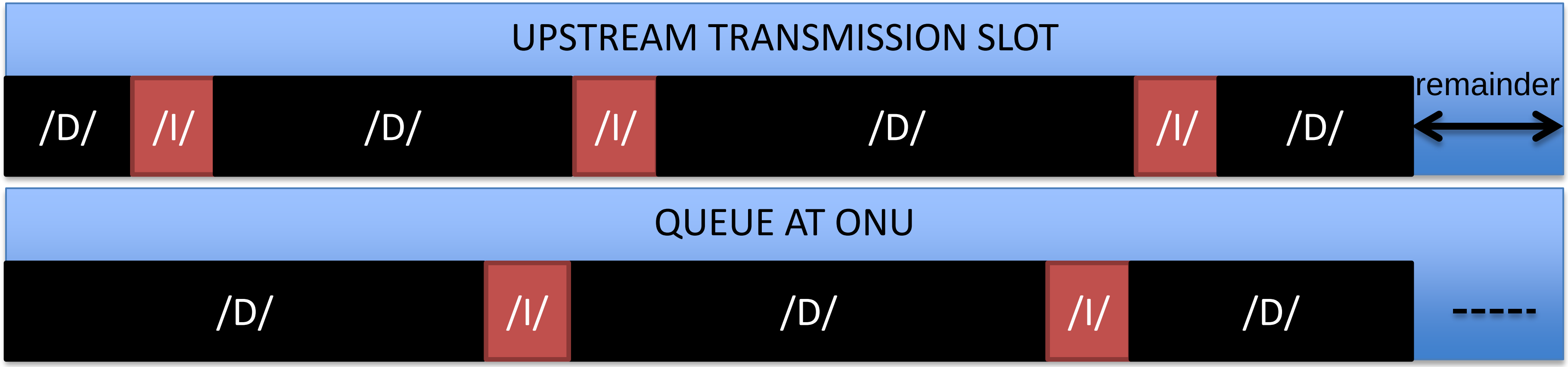
Upstream slot model

- To simplify discussion, assume that we have a simple model of an upstream transmission slot (EPON slot)
  - Length: slot_length
  - Start time: slot_start



Upstream slot remainder – limitations

- In Ethernet, frames are non-fragmentable
  - If a frame does not fit completely into the allocated slot, such a frame is delayed until the next transmission grant opens for the CNU.
- Frame reordering is not allowed by 802
  - Within a single "conversation*", ONU cannot transmit smaller frame from further down the queue - that affects operation of upper layer protocols
  - It is still possible to prioritize between different "conversations"



next packet for transmission is greater than the available slot size

* See IEEE Std 802.1X, 3.8, for a definition of a "conversation"

Upstream slot remainder - Scenarios

- Conditions for occurrence of slot remainder:

  Scenario A: ONU reports queue occupation without frame delineation , e.g., reporting 0xFF-FF when queue occupancy crosses the last threshold;

  Scenario B: OLT grants less bandwidth than ONU requested and queue delineation boundaries are not observed, e.g., ONU reported 2750 octets in queue, OLT grants only 2000 octets.

  Scenario C: More high priority frames arrive at the ONU queue between the time of reporting and the start of upstream slot. In ONUs that allow transmission of unreported frames, this changes frame delineation boundaries and leads to slot remainder

- Any combinations of these scenarios are also possible
- In any case, ONU may potentially not manage to fill in the granted slot completely due to mismatch between frame delineation boundaries in the ONU queue and granted slot size
  - This leads to slot remainder and waste of upstream bandwidth

Upstream slot remainder – Scenario A

- Typical scenario leading to upstream slot remainder (A):
  - When ONU queue contains data in excess of the last report threshold, ONU typically reports value of 0xFF-FF back to OLT.
  - Such reported value contains no information about frame delineation (where the last frame ends)
  - OLT cannot figure out where the last complete Ethernet frame ends due to lack of delineation information
  - The grant size is selected by OLT using local, vendor-specific policies and QoS enforcement mechanisms
  - ONU tries to fill in the granted slot size, but it typically does not coincide with frame boundaries within ONU queue
- This problem typically occurs under heavy load, where ONU has more data to report than REPORT MPCPDU can support
- It can be mitigated by ONU reporting multiple queue thresholds and OLT granting only on one of the requested thresholds

Upstream slot remainder – Scenario B

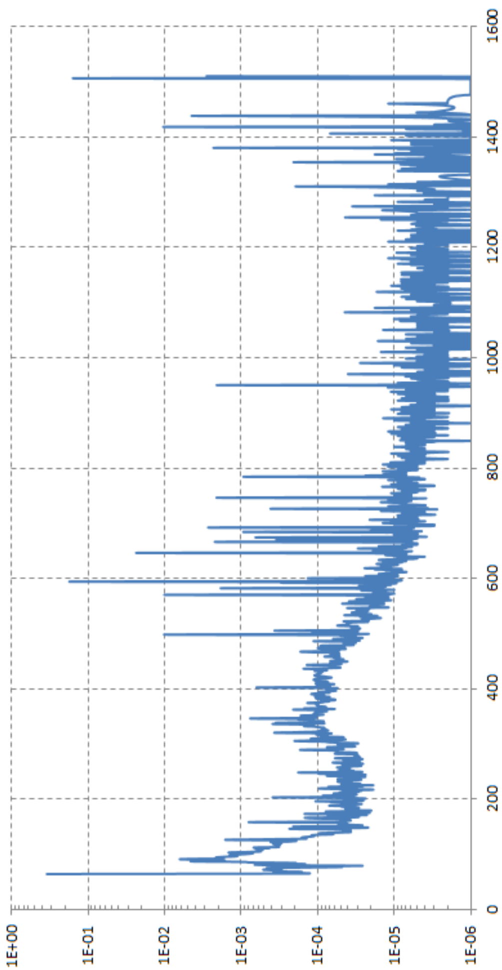
- Typical scenario leading to upstream slot remainder (B):
  - In this scenario, ONU reports bandwidth with frame delineation boundaries, i.e. reported value is smaller than 0xFF-FF;
  - Under heavily load conditions, OLT may not be able to grant upstream slot equal to any reported ONU bandwidth demand
  - OLT will have to grant smaller upstream slot, losing in this way frame delineation boundaries at ONU
- This problem typically occurs under upstream heavy load, where OLT cannot grant ONU the bandwidth it reported
- This problem can be mitigated by ONU reporting multiple queue thresholds and OLT granting only on one of the requested thresholds

Upstream slot remainder – Scenario C

- Typical scenario leading to upstream slot remainder (C):
  - In this scenario, ONU reports bandwidth with frame delineation boundaries, i.e. reported value is smaller than 0xFF-FF;
  - OLT grants according to bandwidth requested by ONU (following one of reported threshold values)
  - ONU implements strict priority policy, where higher priority frame arriving at ONU queue between the time of last report and next grant displace some lower priority frames.
  - Higher priority frames are sent upstream (even though some were not reported), invalidating reported frames boundaries.
- This problem can occur under <u>any upstream load</u> scenario
- To mitigate this problem, ONU should send upstream only previously reported frames, no matter what frame priority

Average upstream slot remainder (I)

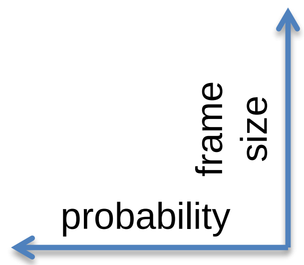
- Average upstream slot remainder size under heavy load conditions can be assessed statistically, if PSD is known

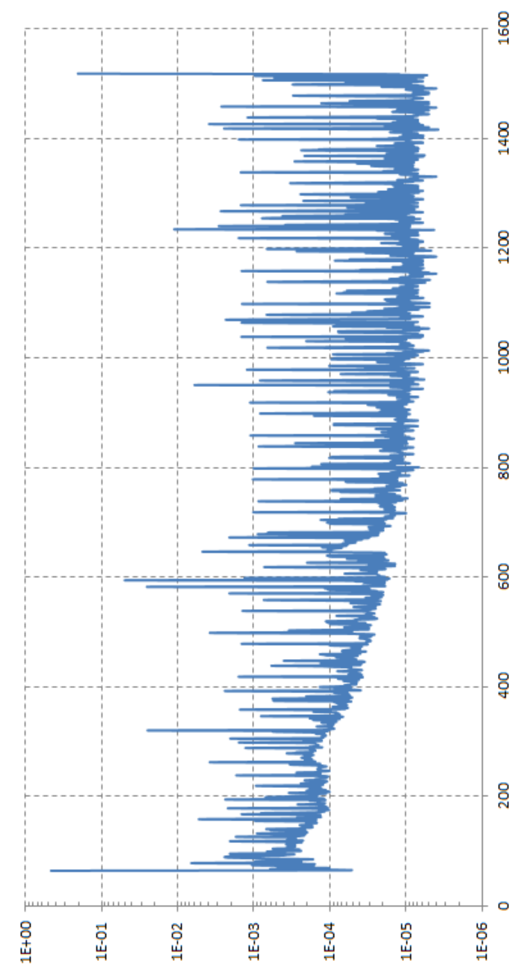


downstream direction
average packet size 673 octets

upstream direction
average packet size 530 octets

Upstream and downstream traffic profiles at the same network point (CMTS), captured in 2001





Average upstream slot remainder (II)

- For a known PSD, the average (expected) size of upstream slot remainder can be calculated*:

$$E(R) = \frac{1}{E(X)} \sum_{r=1}^{M-1} r\left[1 - F_X(r)\right],$$

where:

F_X(r) denotes a cumulative distribution function for packet size **r**

M denotes the maximum frame size (1518 octets)

- Value for traces shown on previous slide
  - ~617 octets for downstream
  - ~597 octets for upstream

* “Supporting differentiated classes of service in Ethernet passive optical networks”, G. Kramer et al., http://www.glenkramer.com/ucdavis/papers/cos_jon.pdf

Coarse granting / reporting (I)

- In 1G-EPON, 1 symbol (TQ) is worth 2 data octets.
- In 10G-EPON, 1 symbol (TQ) is worth 20 data octets
- Granting and reporting takes place in units of TQ, defining resolution for these processes:
  - ONU rounds up bandwidth demand when reporting, OLT has no way to calculate actual bandwidth demand
  - Frame delineation problems are likely to occur in this scenario, even if all sources of upstream slot remainder are accounted for
  - Average slot remainder is smaller here: 1 octet for 1G-EPON and 10 octets for 10G-EPON (on average, half of TQ size in octets)
- In 10G-EPON, the problem is further exacerbated by mandatory, stream-based FEC (255 octets long)
  - Upstream granting can be done only in multiples of FEC words

Coarse granting / reporting (II)

- This problem would become even more serious in EPoC, if the basic granting symbol were equal to a whole OFDMA symbol
  - Past presentations places ODFMA symbol size anywhere between 10us and 1ms, resulting in 625 - 62,500 TQ symbol sizes
  - CLT might chose to grant only complete OFDMA symbols to CNUs
  - This would mean that even if CNU requests less bandwidth than a single OFDMA symbol, OLT would still grant multiples of complete symbols, resulting in a large bandwidth remainder.
- The ability for multiple CNUs to share a single OFDMA symbol in upstream is crucial for increased efficiency
  - 1D or 2D scheduling needs further discussion, to properly assess the impact of using time-based or time-and-frequency based scheduling on upstream efficiency

Conclusions

- Upstream efficiency is heavily impacted by minimum symbol size
  - Sharing a single symbol among multiple CNUs is a MUST for EPoC to avoid substantial bandwidth loss at MAC level
  - This is the only item that is really within the control of the P802.3bn Task Force
- Upstream scheduling, granting and reporting has also substantial impact on efficiency
  - These items are outside the scope of P802.3bn TF and will rely on vendor-specific solutions.
  - Large part might be further specified in future 1904.2 (?) project, in conjunction with CableLabs
- Commercial EPON systems use various ways of improving upstream efficiency even today:
  - Vendor-specific queue service, scheduling and polling policies operating at both the OLT and ONU

Looking forward (I) …

- PHY layer efficiency is only one part of the overall EPoC system efficiency. Let’s keep in mind impact of individual EPoC stack elements on the system level efficiency:
  - All pieces of the puzzle must fall together right
  - FEC code and encoding / interleaving function selection impacts data rate adaptation design and restricts scheduling mechanisms
  - All these factors must be examined together (and not separately)
  - Some of these factors remain outside the scope of P802.3bn TF (scheduling, granting etc.)

Looking forward (II) …

- To move forward, let's build the first PHY stack model, examine its efficiency and try to improve it
  - Trying to get it right and perfect at once almost never works
  - Discussions without a baseline stack model take forever and lead to heated discussions without clearly formulated arguments
  - We need a baseline model out of November to start working on contributions for individual functions for January 2013 meeting
  - It can be done at a conceptual level during the meeting, adopted and then detailed proposals for individual sublayers could be delivered for January 2013 meeting

More reading on EPON efficiency …

- "Supporting differentiated classes of service in Ethernet passive optical networks", by G. Kramer, et al.,
- "Cyclic Polling-Based Dynamic Bandwidth Allocation for Differentiated Classes of Service", by Su-il Choi et al.,
- "Efficient resource allocation with service guarantees in passive optical networks", by T. Orphanoudakis et al.,
- "Ethernet PONs - A survey of dynamic bandwidth allocation (DBA) algorithms" by Michael P. McGarry et al.,
- "IPACT - A Dynamic Bandwidth Distribution Scheme in an Optical Access Network" by G. Kramer et al.,
- "Ethernet Passive Optical Networks" by G. Kramer
- "Delivering Carrier Ethernet: Extending Ethernet Beyond the LAN" by Abdul Kasim et al.,
- "On Efficiency of Ethernet Passive Optical Networks (EPONs)" by M. Hajduczenia et al.,
- "Performance of 10G-EPON" by Rajesh Roy et al.,

Bringing you Closer

Thanks!