GLID as Envelope ID – a Deeper Look

Glen Kramer, Broadcom

Name vs. Essence

NG·EPON

- **"X"** is a name of entity that can be independently scheduled by the OLT and thus allow independent QoS and traffic isolation from other such entities.
  - The number of entities of **"X"** provisioned per PON is a service quality and policy issue, thus an operator's decision

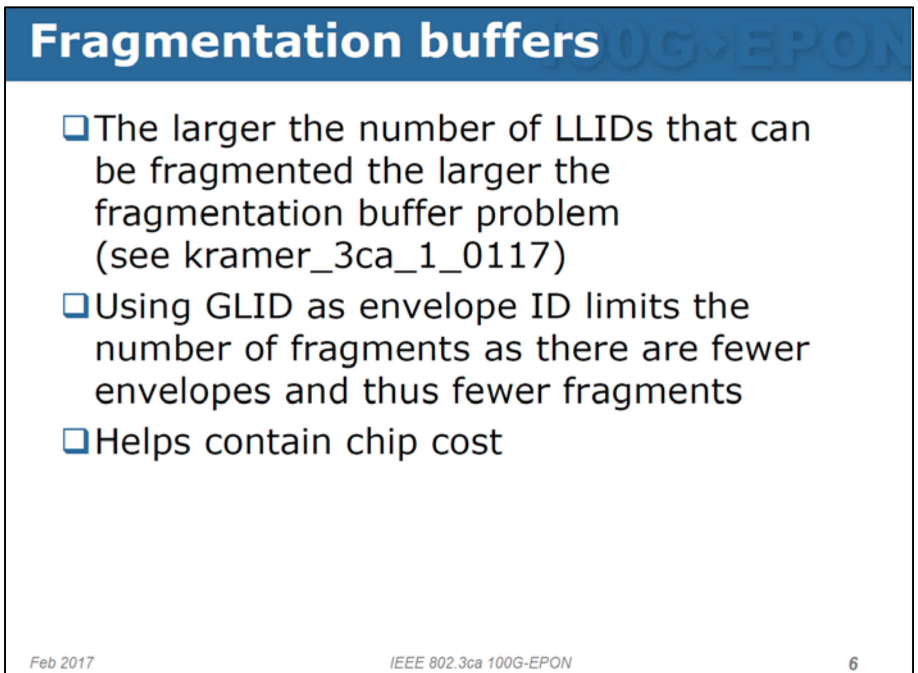
- **"Y"** is a name of entity that identifies segregation of network traffic into individual logical connections, flows, streams, sessions, etc. **"Y"** entities <u>are not independently scheduled</u> and may not allow independent QoS or isolation from other such entities.

- What **"X"** and **"Y"** are actually called

	In ITU-T PON	In IEEE PON	In remein_3ca_3_0317
"X"	Alloc-ID (DS only)	LLID	GLID
"Y"	(X)GEM Port-ID	VLAN (VID)	LLID

What is the problem?

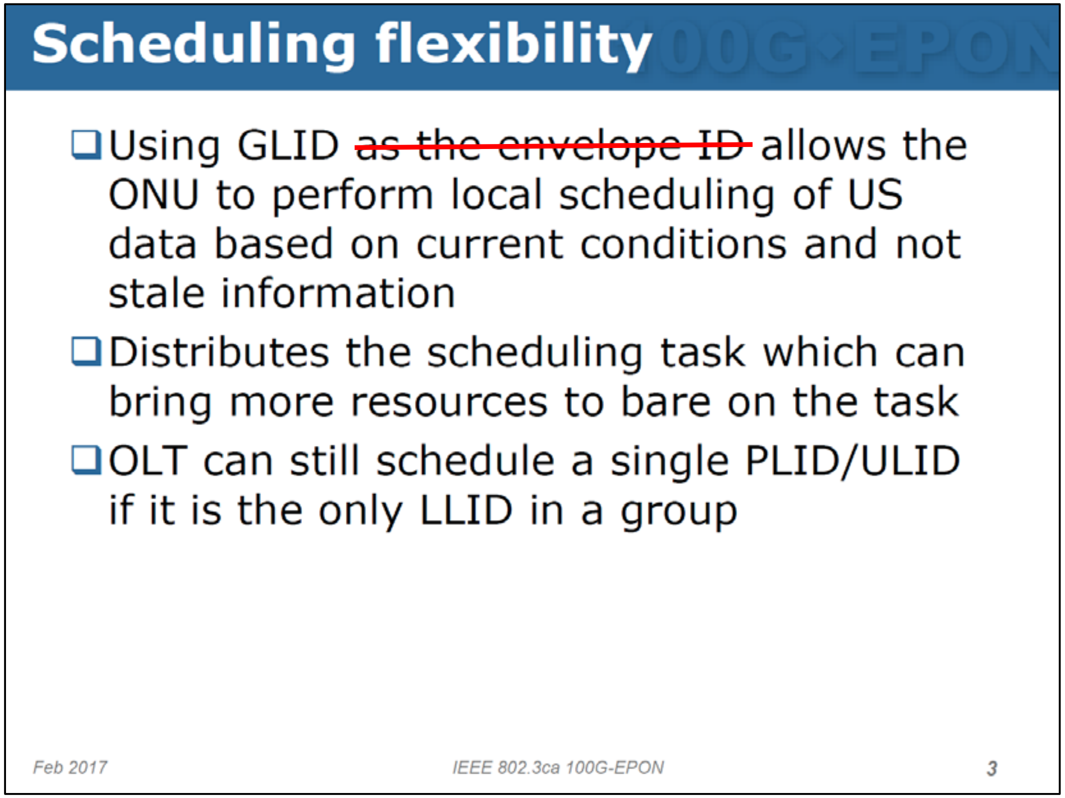
- The number of LLIDs provisioned in EPON is a service quality and policy issue.
  - This requirement comes first and is external to PON
  - LLIDs are provisioned if an when necessary. If the number of LLIDs can be reduced, then they should not have been provisioned in the first place.
- The reassembly buffer size is an internal design issue. It should not restrict the number of LLIDs.
- Remein_3ca_3_0317 does not solve the problem. It just suggests that
  1. The number of schedulable entities **“X”** be reduced.
  2. A tag be added to identify entities **“Y”** (in addition to VLAN tags typically used for this).



- The number of schedulable entities **“X”** cannot be reduced, because it is determined by QoS and policy requirements.
- The “GLID as Envelope ID” proposal doesn’t solve anything. All it does, it renames “LLID” into “GLID”, i.e., instead of scheduling ***N*** LLIDs it will now schedule ***N*** GLIDs.

What is the problem?

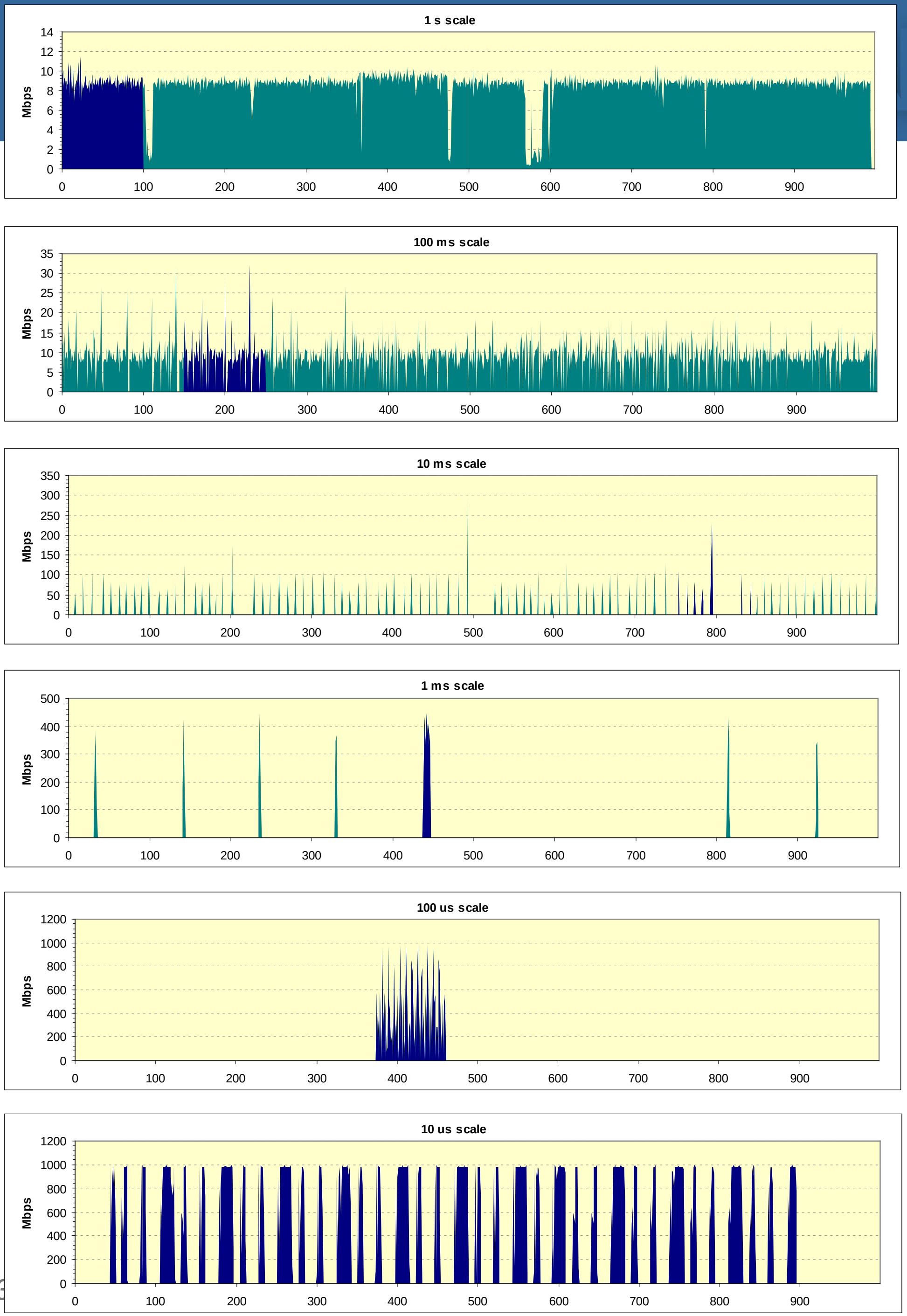
- Bullets 1 and 2 just show the general motivation of using GLID for upstream scheduling. <u>This has nothing to do with using GLID as the envelope ID!</u>
- Bullet 3: If the OLT still needs to schedule single PLID/ULID, then the proposed scheme is equivalent to what is accepted in D0.2 already.



- Note that this scheme requires every ULID to always be either scheduled as part of a group or individually (i.e., a group of one)
- It does not allow the OLT to make this decision dynamically at run-time (e.g., based on current load, number of active LLIDs, or user behavior).

Traffic burstiness

- Traffic is bursty at many timescales – microseconds to hours.
- Even single flows are bursty on sub-RTT timescales (see caida.org/workshops/isma/0411/slides/dovrolis.ppt)
- This means that if a frame belongs to a flow A, the next received frame is much more likely to belong to the same flow A, and not to another flow B.

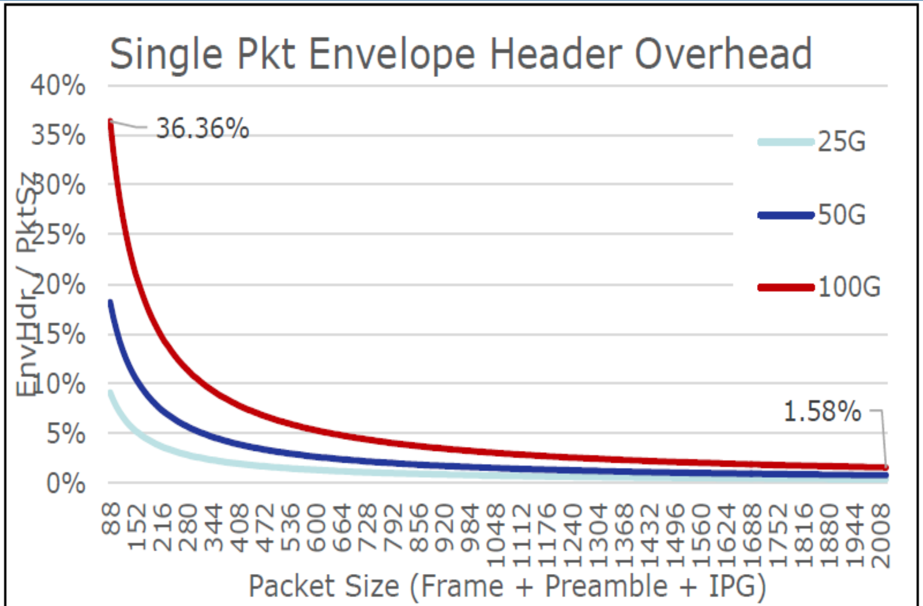


Envelope Header Overhead

- This chart shows Envelope Header Overhead under the following conditions:
  1. All packets are of same size (starting with all 64-byte packets)
  2. Every packet goes to a different LLID (every envelope contains a single packet)
  3. Every packet to 100G ONU is split over 4 wavelength

- These are not realistic conditions and they cannot happen all at once.

- Packets usually arrive in bursts, so there are usually multiple packets that can go in each envelope.



- If multiple LLIDs all have a single packet to transmit, then up to 4 LLIDs can transmit in parallel on 4 wavelengths – no need ever to split across wavelengths.

- The 36% overhead is not possible, unless OLT deliberately tries to self-sabotage.
  - The same self-sabotage is possible in XG-PON, because XGEM header is 8 bytes and the minimal XGEM Payload Length is 1 byte (but it is rounded up to 4 bytes). So, the worst case overhead is 91.7%. But it has never been considered a problem because devices do not deliberately sabotage their performance.

Problems with this proposal

- "PLID/ULID is a member of only one GLID"
  - To allow QoS, specifically CIR+EIR bandwidth control per ULID, ULIDs may need to be members of more than one group (see example on the next page)
  - Restricting ULID to membership in only one group is a major limitation, which makes GLIDs quite useless.

Realization of GLID as EnvID

- Restrict LLID in Envelope header to be only GLID
- Allow an envelope to carry multiple PLID / ULIDs
- Utilize the same preamble replacement as is done for 10G-EPON and 1G-EPON to carry PLID/ULID
  - Envelope header describes the transport envelope
  - PLID/ULID in preamble identifies the MAC of the frame
- Some restrictions apply
  - PLID/ULID is a member of only one GLID
  - May want to consider allowing one set of GLIDs for US use and a different set for DS use (in this case a PLID/ULID would be a member of one GLID per direction)
    - Would only need 3 GLIDs in DS in a 100G-EPON mixed system;
      1 for 25G ONUs (Ch 0),
      1 for 50G ONUs (Ch0 & 1), and
      1 for 100G ONUs (4 Chs)

Feb 2017 IEEE 802.3ca 100G-EPON 8

- "Would only need 3 GLIDs in DS in a 100G-EPON mixed system;
  - 1 for 25G ONUs (Ch 0),
  - 1 for 50G ONUs (Ch0 & 1), and
  - 1 for 100G ONUs (4 Chs)"

- First, this would preclude single-copy broadcast or multicast involving different ONUs.
- Second, the receiving MPRS needs to filter/demux based on PLID/ULID values to find the right destination MAC. GLID values in the envelope header are useless!

GLID Use Example (1/2)

- Consider an MDU ONU with **N** ULIDs (**N** can be very large, but to simplify the example, we take **N**=4) with the following targets for CIR/EIR

ULID	CIR (Mbps)	EIR (weights)	Explanation
A	500	0	Guaranteed BW of 500 Mbps, no extra bandwidth
B	500	1	Guaranteed BW of 500 Mbps, plus extra bandwidth may be given if available
C	1000	1	Guaranteed BW of 1 Gbps, the extra bandwidth is the same as for ULID B
D	2000	3	Guaranteed BW of 2 Gbps, the extra bandwidth is 3x of what ULIDs B or C get.

- To grant these ULIDs, we need two GLIDs:
  - GLID1 = {A,B,C,D} – grants bandwidth to satisfy CIR
  - GLID2 = {B,C,D} – grants bandwidth for EIR

GLID1	GLID2
Assigns following weights A:1, B:1, C:2, D:4	Assigns following weights B:1, C:1, D:3
Grant length ***L*** is dependent on the time interval ***T*** since the last grant to GLID1: $L = \sum CIR \times T = 4Gbps \times T$ If T = 2 ms, L = 125,000 EQ If T = 4 ms, L = 250,000 EQ The grant bandwidth is distributed among ULIDs according to CIR weights	• Grant length is determined based on overall EPON load and time available to next time-sensitive (pre-scheduled) grant. • Under the high load conditions, GLID2 may not be granted for a long time. • When GLID2 is granted, bandwidth is always distributed among ULIDs according to EIR weights

GLID Use Example (2/2)

- Now, consider a GATE arrived with these two grants

GLID1	GLID2
Grant Length = 200,000 EQ	Grant Length = 120,000 EQ

- The following is the correct allocation of envelope sizes per each ULID:

ULID	From GLID1 (EQ)	From GLID2 (EQ)	Total Envelope Length
A	25,000	n/a	25,000
B	25,000	24,000	49,000
C	50,000	24,000	74,000
D	100,000	72,000	172,000

- If a ULID can only belong to one group (i.e., to one GLID), such CIR+EIR allocation will not be possible.

Thank You