Contribution to
The IEEE 802.3 100G EPON Task Force Meeting, July 2016

Bonding with FEC Codewords

Frank Effenberger, Duane Remein
Fixed Access Network Research
Futurewei Technologies, Huawei R&D USA

www.huawei.com

HUAWEI TECHNOLOGIES CO., LTD.

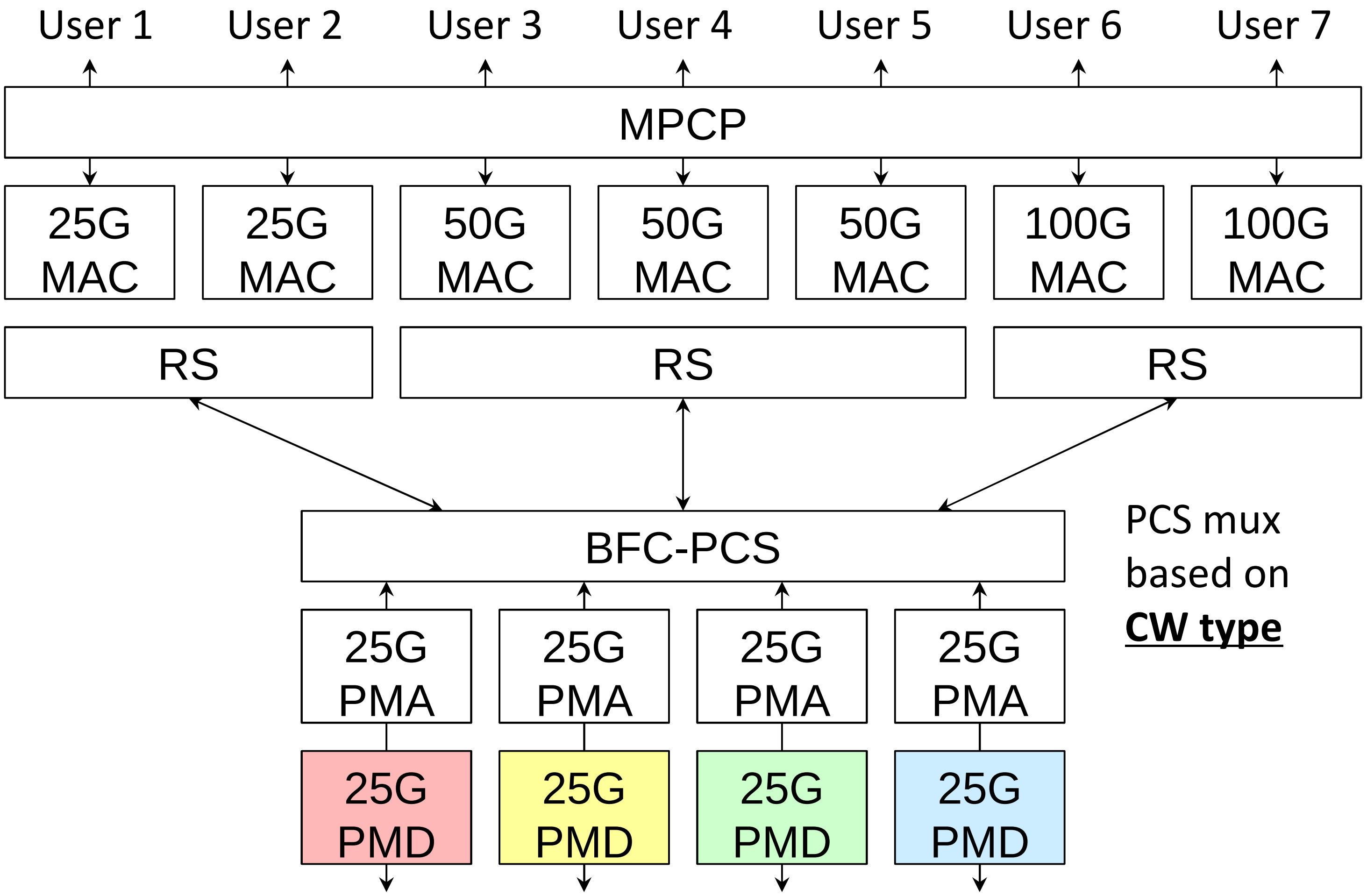
HUAWEI

Introduction

- From early in the project, there has been a stated desire to support "100G MAC rates"
- Previous analysis (Kramer, March'16) showed that bonding at the lower layer broke the DBA
- But that was under the assumption that whole frames would be sent on each lane

- This presentation explores the possibility of bonding at the FEC codeword level (aka Striping)
- We assume that the channels can be "aligned" at the CW level, so CW ordering is maintained
  - Some sort of markers could be used
  - This is how 100G Ethernet works

OLT architecture

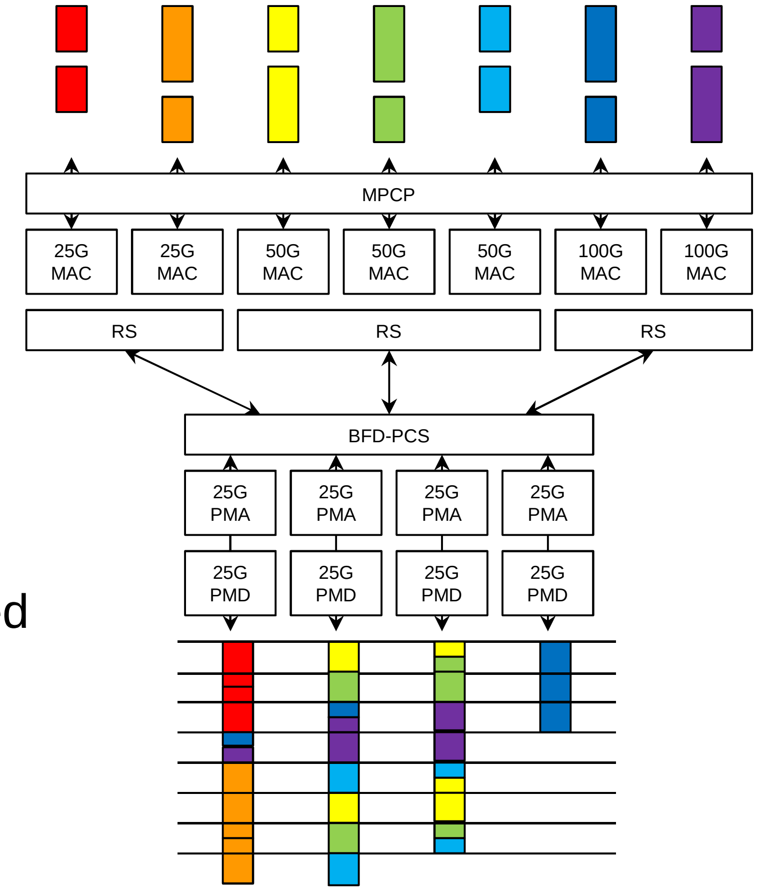


RS mux based on LLID

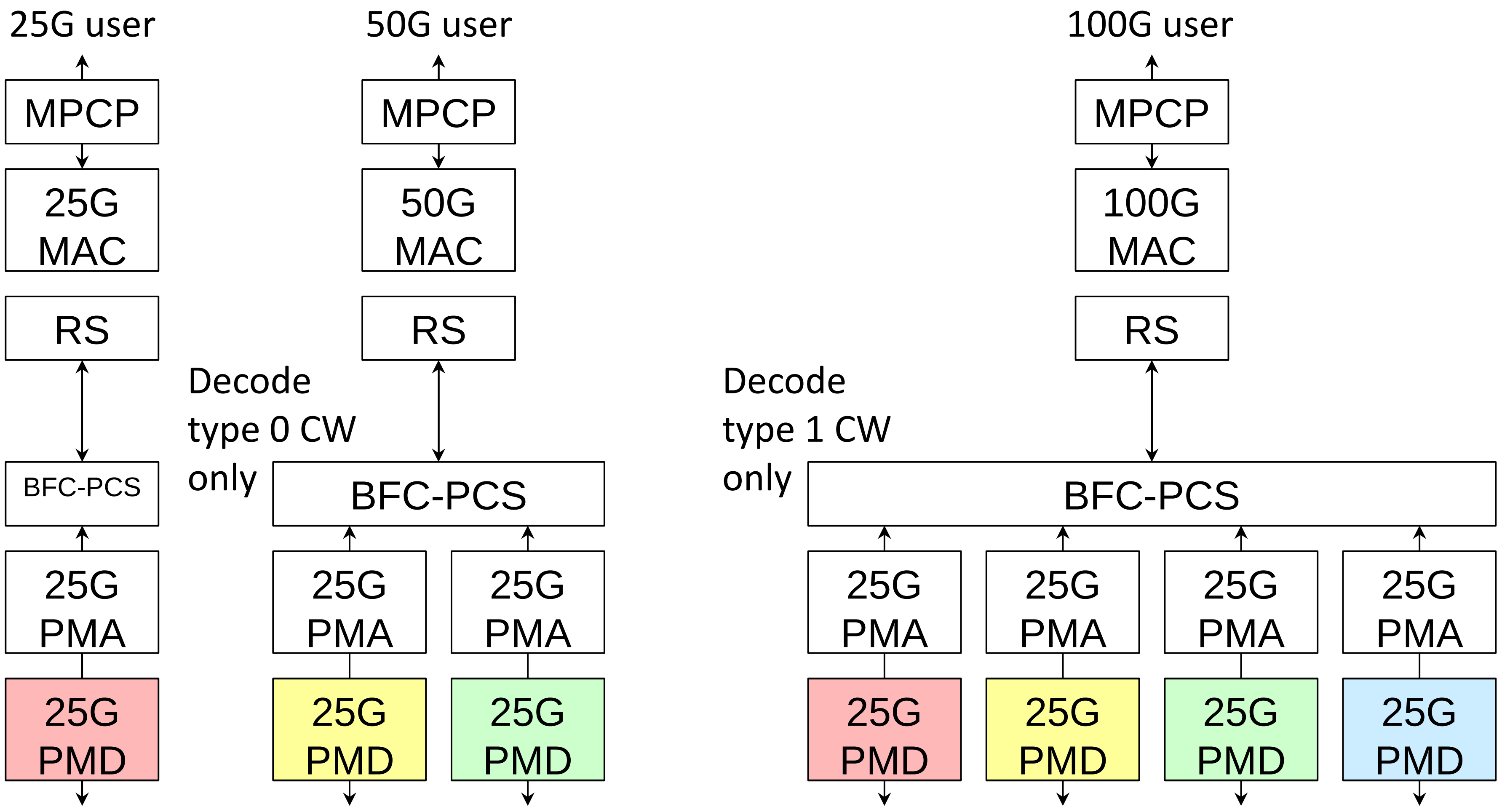
PCS mux based on **<u>CW type</u>**

CW multiplexing example

- How this works
- MPCP times transmission of frames to the MACs
- RS layer marks frames with LLID
- 25G frames seize channel 1
- 50G frames seize channels 2+3
  - FEC CW's are round robin assigned
- 100G frames claim remaining channels, using RR assignment
- 25G and 50G FEC CW are type 0
- 100G FEC CW are type 1



ONU architectures

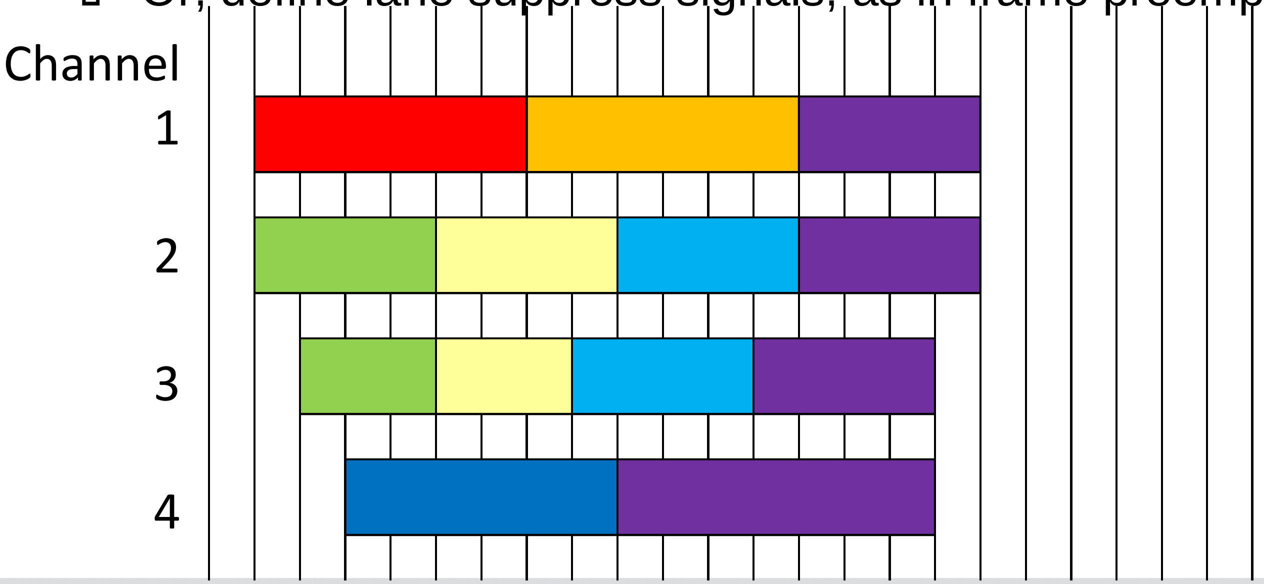


A brief aside: assignment of channels

- Above discussion assumes that the 25G and 50G do not share a channel (the 1, 2-3, 1-4 plan)
- This results in the following BW constraints
  - 25G load < 25G
  - 50G load < 50G
  - 100G load + 50G load + 25G load < 100G
- The previously discussed arrangement (25 and 50 share a channel, the 1, 1-2, 1-4 plan) gives the following constraints
  - 25G load < 25G
  - 50G load + 25G load < 50G
  - 100G load + 50G load + 25G load < 100G
  - Clearly, this is more constraining – less flexible in BW usage
- Relevant to BFC, the 1, 2-3, 1-4 plan requires only two codeword types
  - The 1, 1-2, 1-4 plan can work, but requires three CW types

Upstream assignment

- Each lane can be granted arbitrarily
- Assuming lane alignment, round robin codeword order maintained
- 100G ONUs need way for MPCP to suppress lanes
  - Could have virtual 25G and 50G MAC ports
  - Sending dummy data on these ports will prevent their use for 100G
  - Or, define lane suppress signals, as in frame preemption



Impact on DBA and fragmentation

- In this scheme, OLT may give a group of overlapping grants to any particular ONU, each grant being a multiple of FEC CW's
- Since frames run over all these grants, there isn't fragmentation loss within that overlapping group
- OLT can choose to make the sum of the grants = a watermark
  - Last frame ends neatly in the last codeword
- Alternatively, if we allow for pre-emption, frame can be broken over multiple adjacent bursts
  - However, frame delay would become indeterminate

Impact on frame delay

- The key is that the MPCP knows what all the MACs are doing
  - Frames are released with sufficient IPG to provide slack
- For 25G and 50G, the actual channel bandwidth is constant
  - Current idle insertion process will create constant frame delay
- For 100G, channel BW varies
  - PCS does "know" what the instantaneous BW is
  - A suitable idle insertion method should be possible
  - Also note that MPCP messages fit within 1 CW (might straddle the boundary), so latency of these important frames is limited
- Further work is needed to develop the insertion method

Thank you

www.huawei.com