

# 40 GbE and 100 GbE PCS Considerations

## Key Questions to be Answered concerning OTN mapping for MLD (CTBI) architecture

Stephen J. Trowbridge  
Alcatel-Lucent

# Supporters

---

- Mark Gustlin (Cisco)
- Med Belhadj (Cortina)
- Pete Anslow (Nortel)
- Martin Carroll (Verizon)
- Gary Nicholl (Cisco)
- George Young (AT&T)
- Frank Chang (Vitesse)

## Key Questions regarding PCS and OTN mapping

---

- The “front runner” proposal for a PCS seems to be a MLD(CTBI) type approach that steals from the IPG to make room for lane alignment markers. This generally will result in a different distribution of IPG (idles) at the Rx MII than at the Tx MII. So what does “transparent” transport of 40/100 GbE over OTN given that the LAN itself is not transparent?
  - o Proposed definition (to be verified with operators): The idle redistribution behavior of a LAN-OTN-LAN span should be no different than that of a single LAN span. (Similar to 10 GbE behavior, but not widely recognized by users)
- Historically, transparent service was understood to apply to the serial Ethernet interface at a given rate (e.g., 10G Base-R and not 10G Base-X). What does transparent service mean if IEEE 802.3ba (initially) defines only parallel interfaces?
  - o What would a future 40/100 GbE serial interface look like? Is this a suitable format to carry over OTN? Does the format have to be different based on whether the LAN is serial or parallel?

## Options presented in earlier meetings for OTN mapping

See [trowbridge\\_02\\_0907.pdf](#) and [trowbridge\\_01\\_1107.pdf](#)

- Option 1: Deskew LAN virtual lanes, decode 64B/66B, reinsert into IPG to reach full MAC rate, re-encode 64B/66B (and transcode to 512B/513B for 40 GbE) to carry across OTN
- Option 2: Deskew LAN virtual lanes, remove lane alignment markers from serialized 64B/66B data, and transport rate reduced bitstream (sans lane alignment markers, - 0.0061% reduction in bandwidth) over OTN (PCS and above)
- Option 3: Deskew LAN virtual lanes, leaving lane alignment markers in place to be re-distributed across lanes of the LAN at the far end OTN egress (MLD and above)
- Option 4: Rather than deleting from IPG to make room for lane markers, LAN lanes run at 0.0061% higher bitrate to make room for lane markers "out of band". Deskew and remove lane markers at OTN ingress. This option will also allow for no lane markers once there is a serial PMD for 40 GbE or 100 GbE.
- Option 5: Rather than deleting from IPG to make room for lane markers, LAN lanes run at 0.0061% higher bitrate to make room for lane markers "out of band". Deskew LAN virtual lanes, leaving lane alignment markers in place to be re-distributed across lanes of the LAN at the far end OTN egress

Introduces more rearrangement of idles in LAN-OTN-LAN connection than in single LAN connection

Not possible if LAN deletes from IPG to make room for lane alignment markers

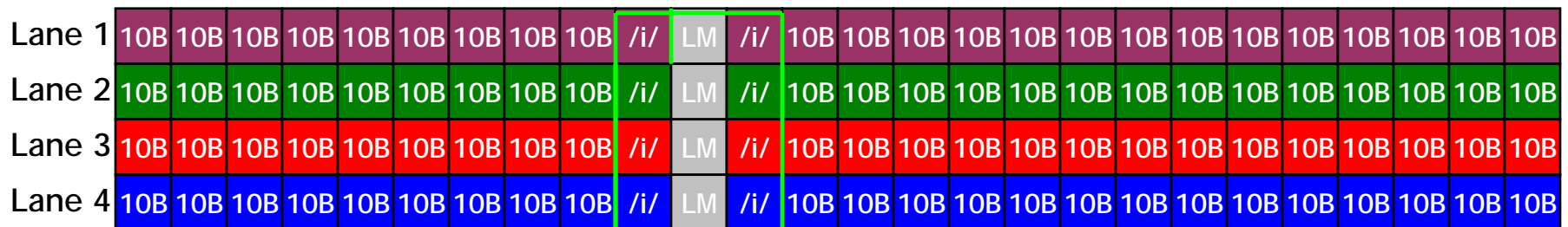
# Lane Striping in 10G Base-X Interfaces



Logical MII order

IPG – 12 byte nominal minimum

≥7 bytes after IPG shrinkage if MII and PHY have different clocks



10B Lane alignment markers replace IPG characters in-place  
Lane alignment markers do not interrupt a packet



Lane Skew

# Lane Striping in 10G Base-X Interfaces

Lane 1	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	LM	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	
Lane 2	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	LM	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B
Lane 3	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	LM	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B
Lane 4	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	LM	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B

Deskew according to lane markers

Lane 1	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	/i/	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B
Lane 2	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	/i/	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B
Lane 3	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	/i/	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B
Lane 4	10B	10B	10B	10B	10B	10B	10B	10B	10B	/i/	/i/	/i/	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B	10B

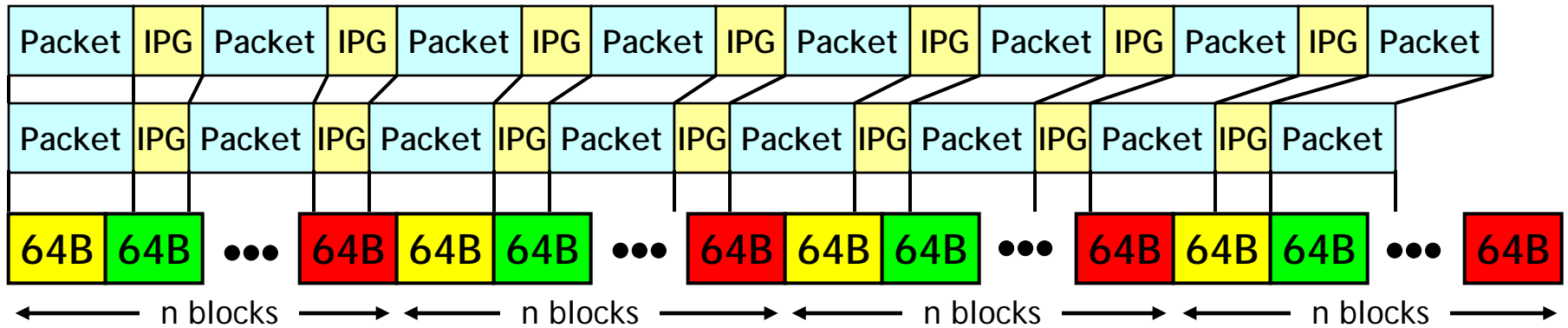
Replace lane markers with idle characters, restoring original IPG

- No rate adaptation anywhere in the process
- IPG may grow or shrink if there is a clock difference between the GMII and the PHY, but never as a result of the lane alignment process (same for all previous Ethernet interfaces)

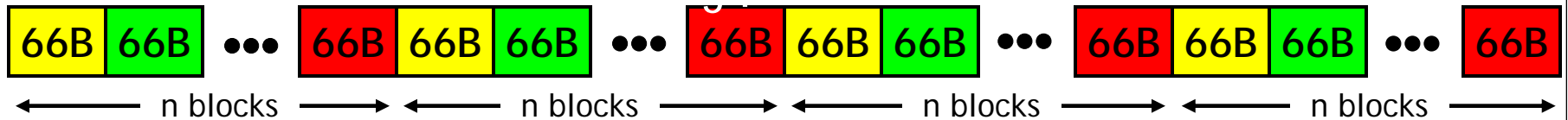
# Lane alignment markers inserted by stealing from IPG

Options 1, 2, 3

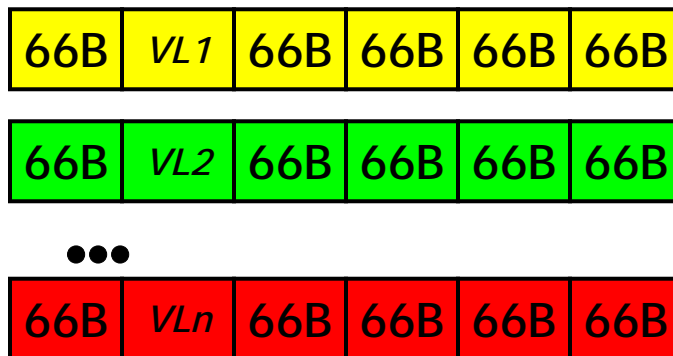
MII: Prior to PCS coding, reduce rate by 0.0061% by deleting from IPG



64B/66B encode:



Inverse multiplex into n (virtual) lanes, add lane alignment markers:

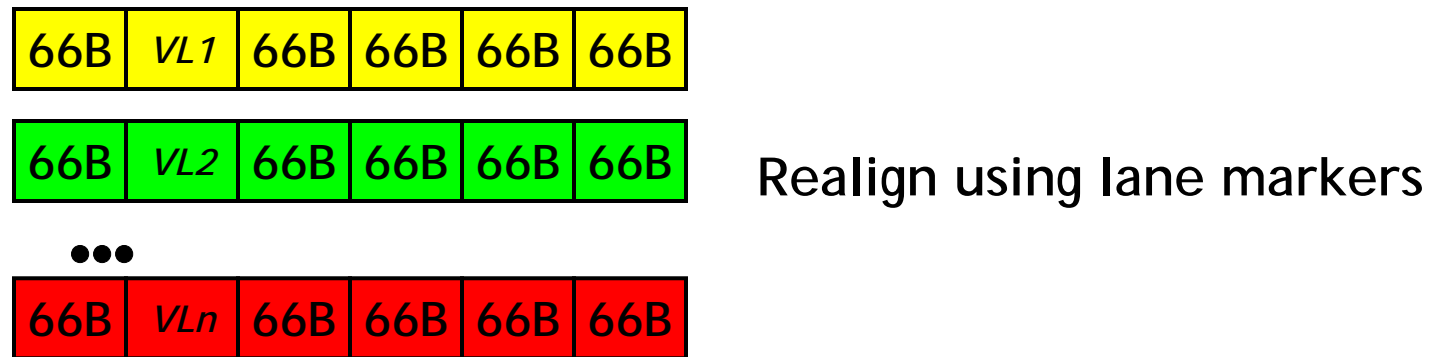
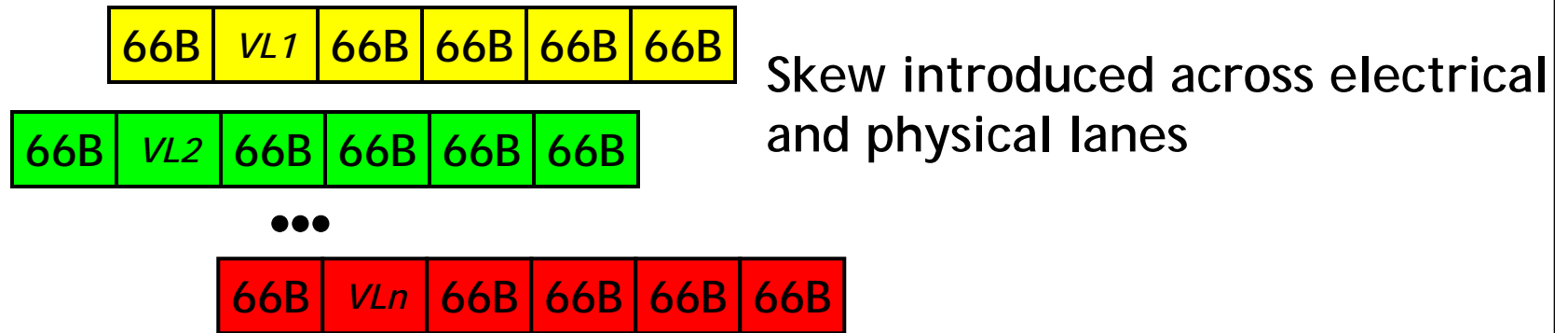


Virtual Lane rate = PCS encoded rate / n

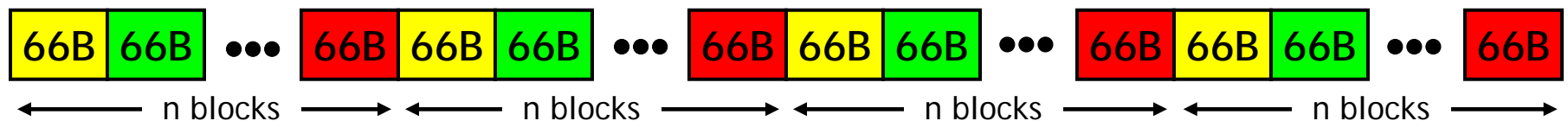
Note: IPG shrinkage must happen before 64B/66B coding in 4(8)-byte granularity because of need to keep packet start in 1<sup>st</sup> (or 5<sup>th</sup>) lane

# Lane alignment markers inserted by stealing from IPG

Options 1, 2, 3



## Serialize by multiplexing virtual lanes and removing lane markers

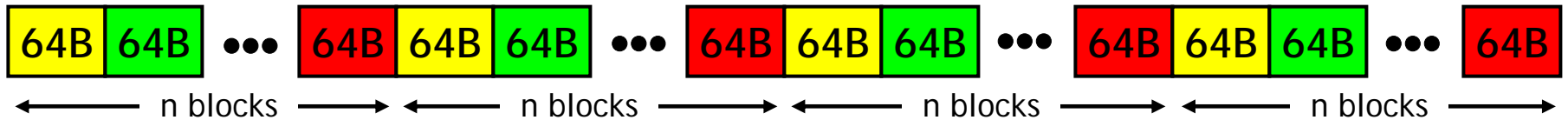




# Lane alignment markers inserted by stealing from IPG

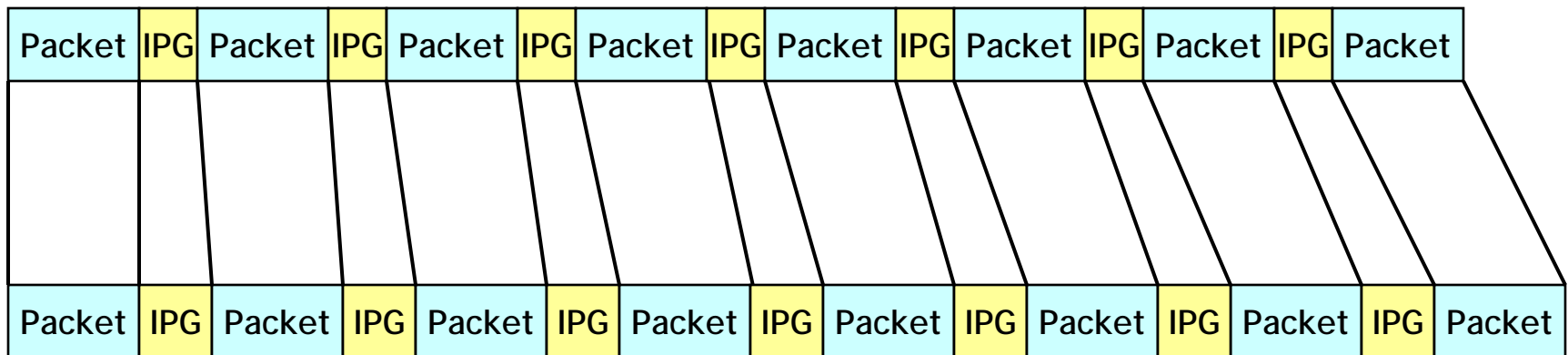
Options 1, 2, 3

## Decode 64B/66B

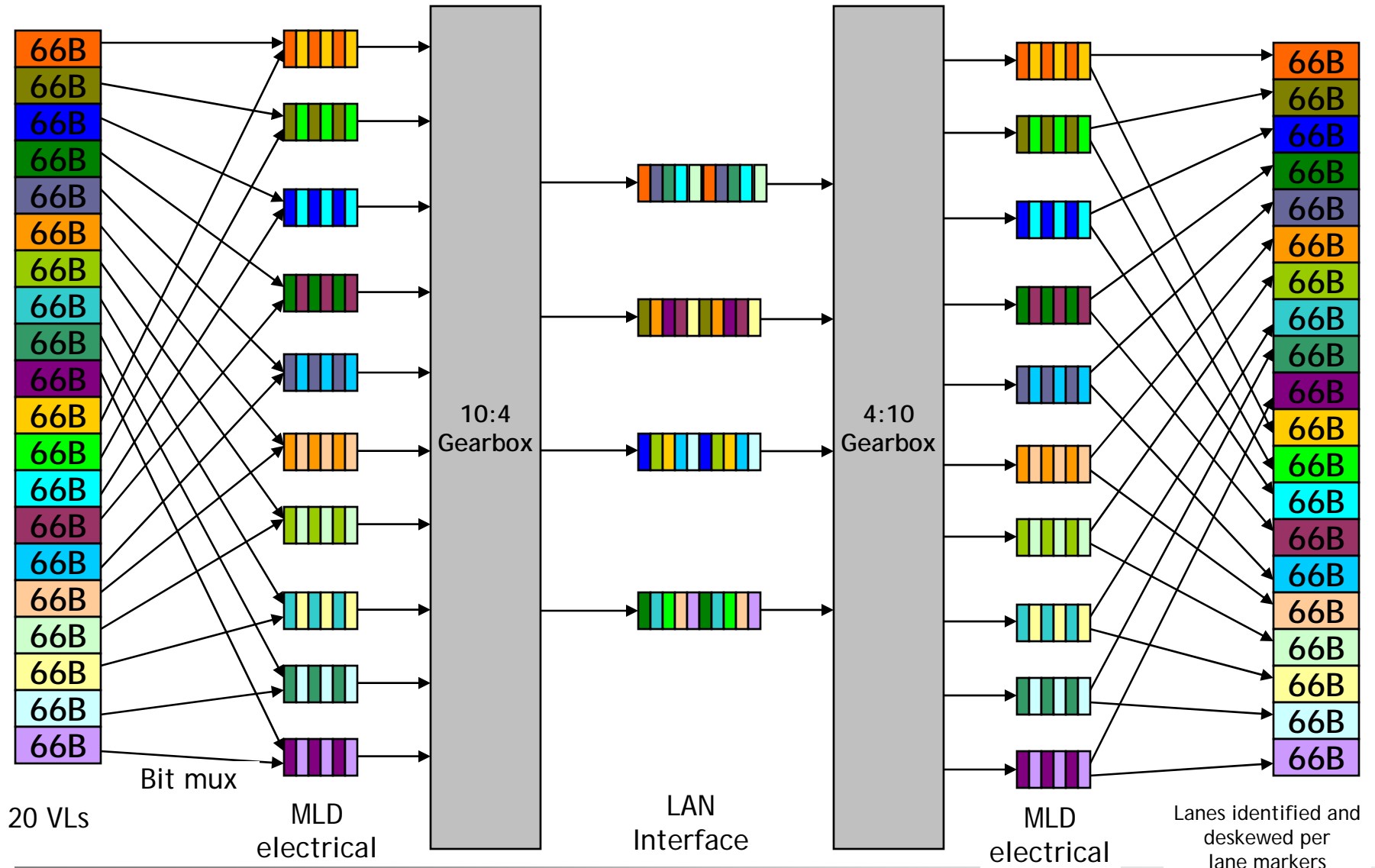


Add 0.0061% by adding to IPG to restore MII rate

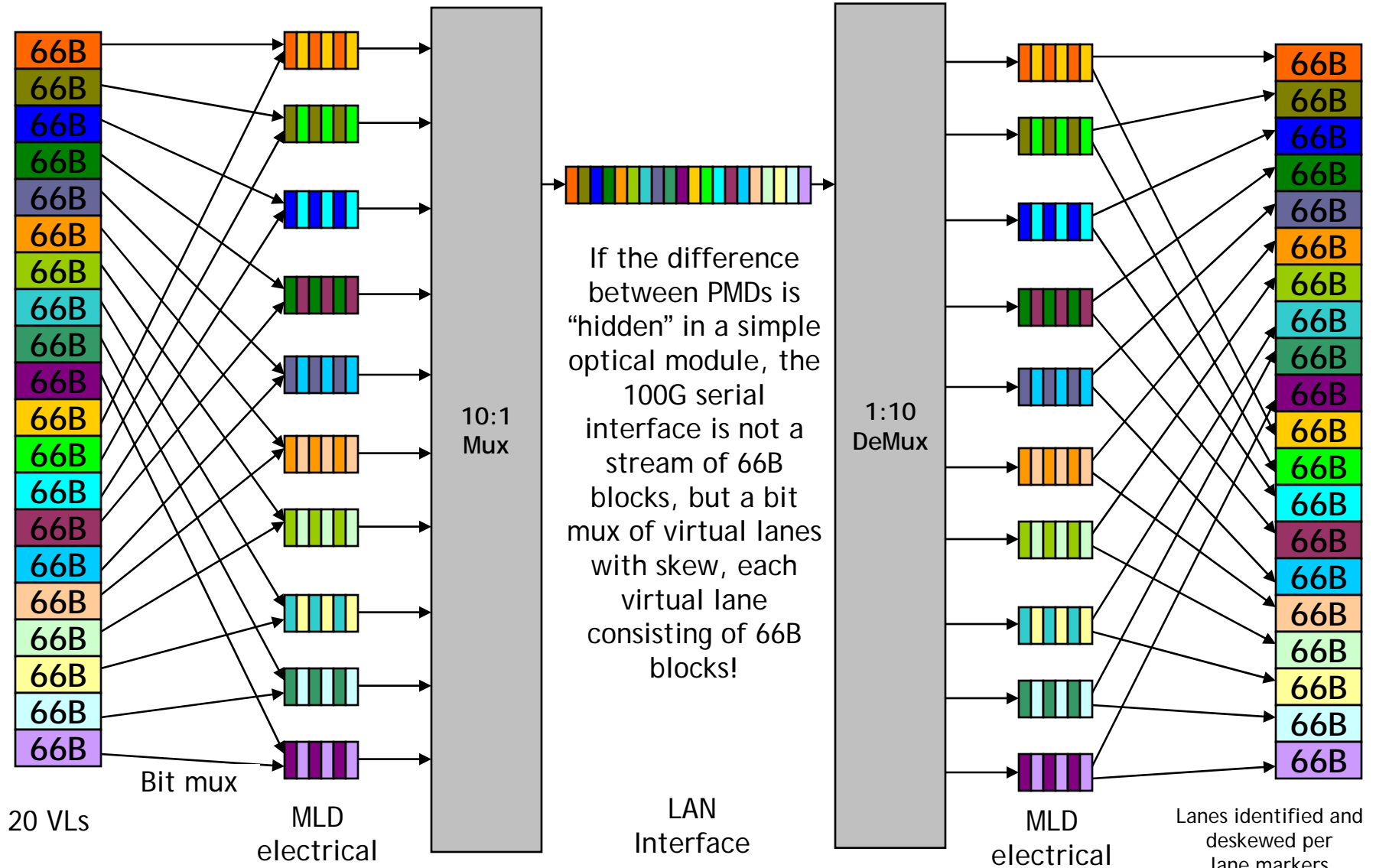
IPG added in 4 (or 8)-byte quanta since both MII and PCS coding rely on packet start in lane 1 (or lane 5) only



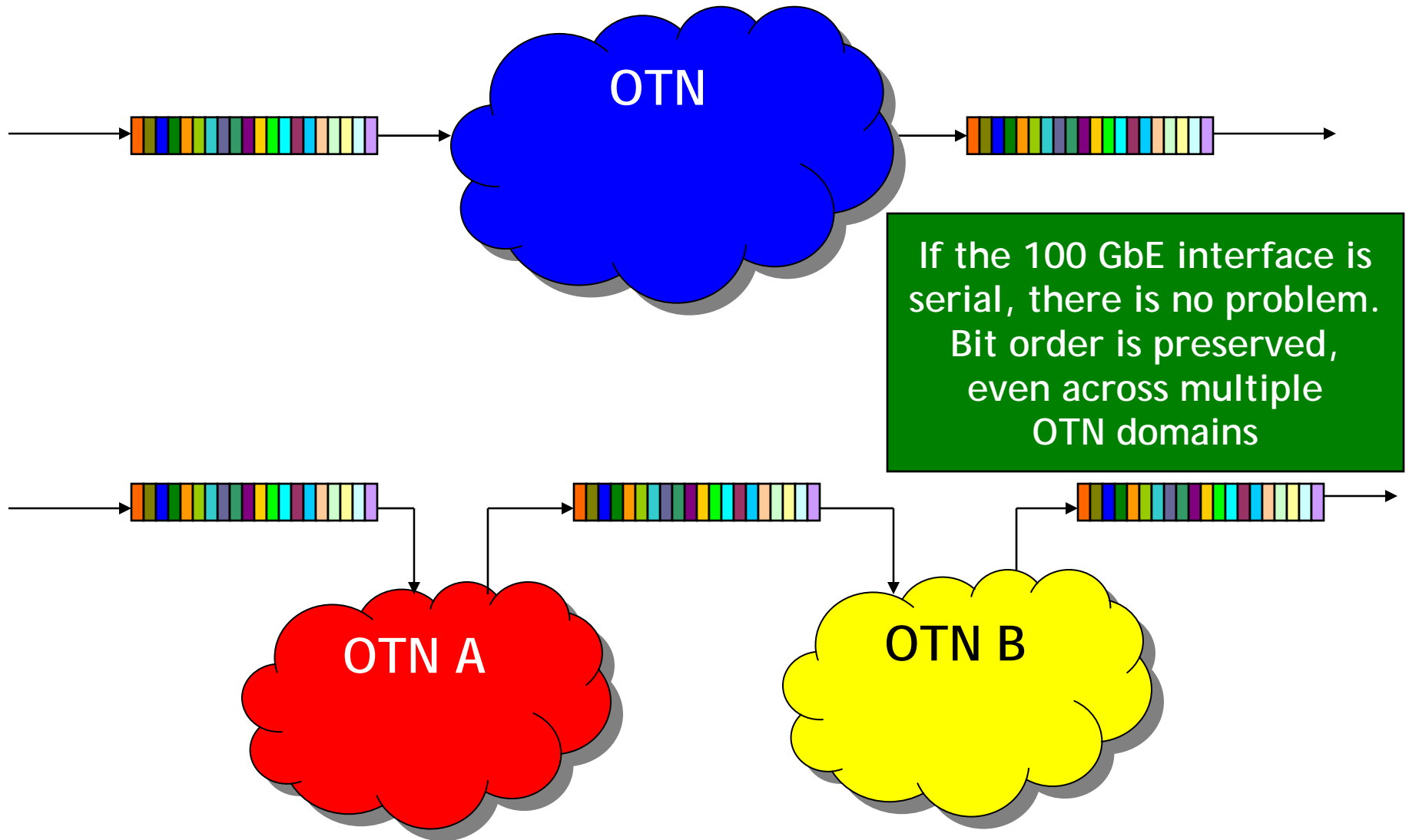
# 100G PCS - MLD(CTBI) into Four-Lane LAN interface



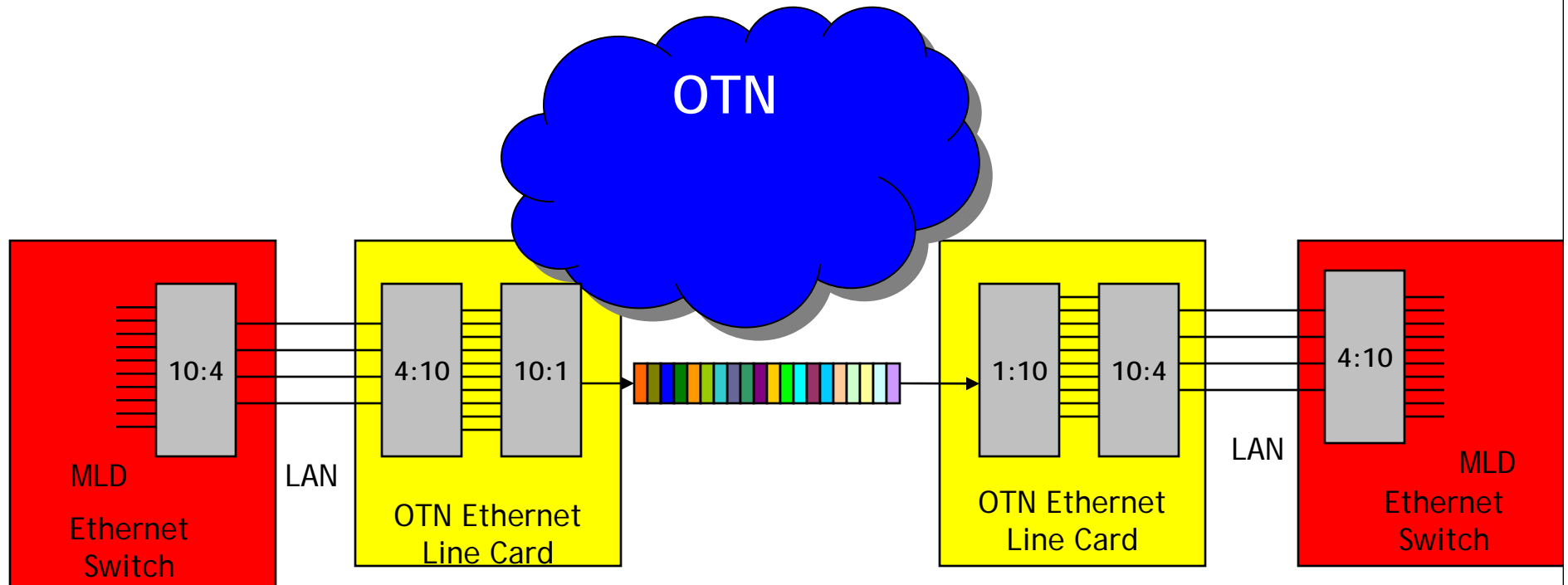
# 100G PCS - What might 100G Serial look like?



Would this kind of 100 GbE Serial Interface format be suitable for an OTN mapping?



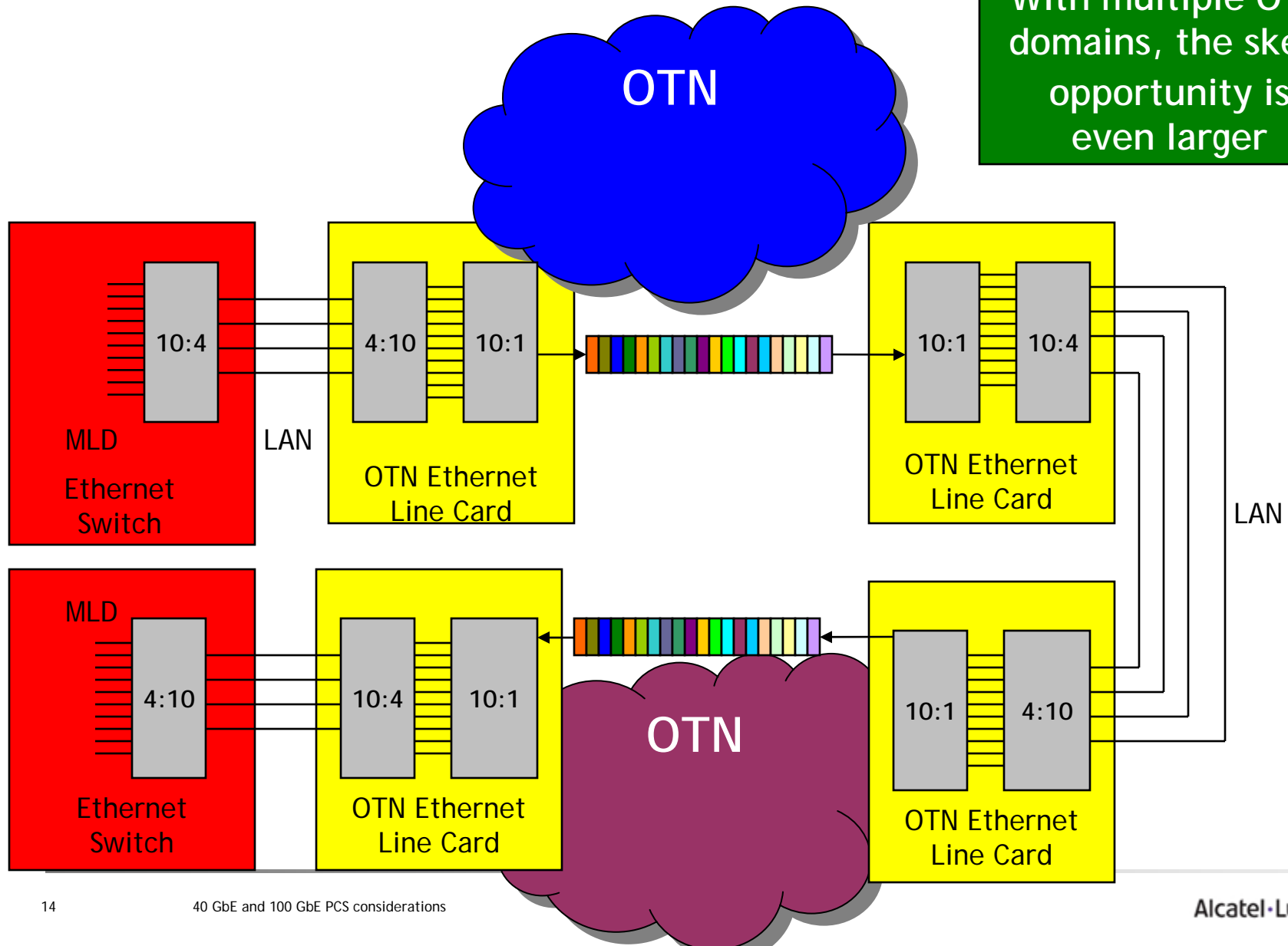
Would this kind of 100 GbE Serial Interface format be suitable for an OTN mapping?



If the LAN is parallel and the serial bit stream is produced by remuxing the bits from the LAN lanes, additional skew opportunity is created between the virtual lanes. A LAN interface across a single OTN domain produces twice the fiber skew and twice the electrical skew (four electrical spans rather than two) of the LAN interface alone

Would this kind of 100 GbE Serial Interface format be suitable for an OTN mapping?

With multiple OTN domains, the skew opportunity is even larger

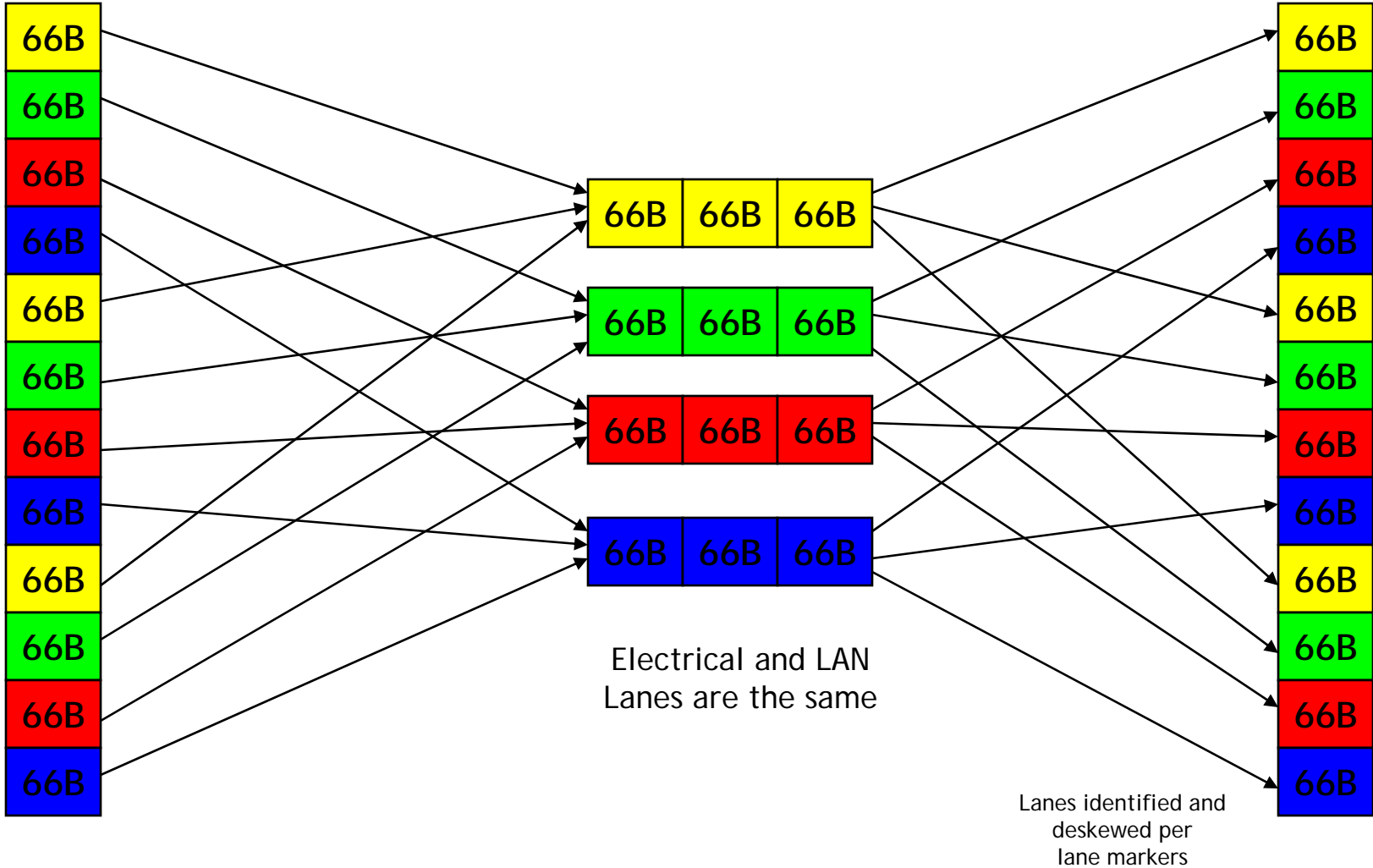


## So what format and architecture is best for 100 GbE over OTN?

---

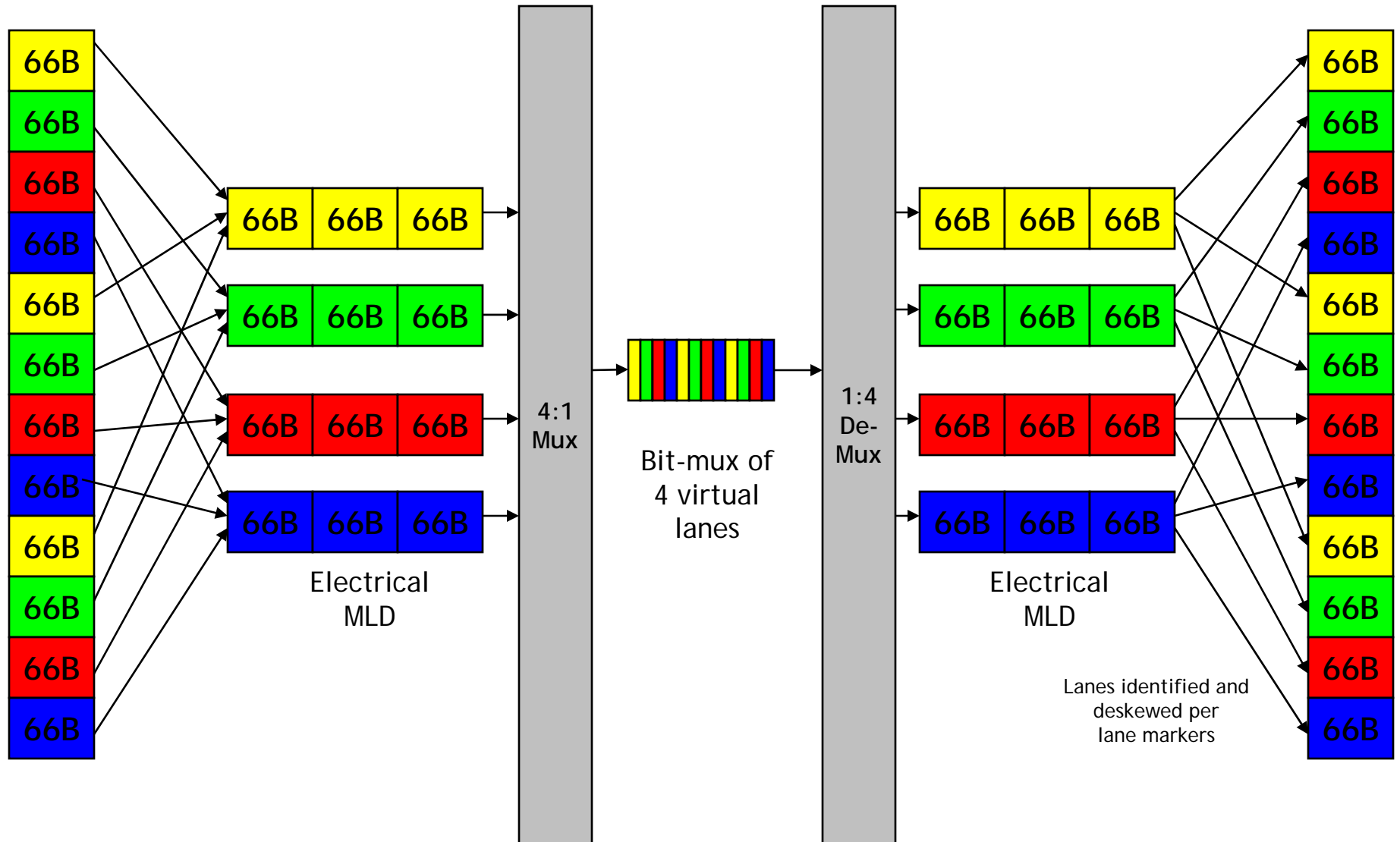
- Option A - Use serial LAN format (bit-mux of 20 virtual lanes) and build the LAN with several times (3-5?) the skew tolerance needed for the longest envisioned parallel LAN interface
- Option B1 - Deskew the LAN virtual lanes at the OTN ingress (requires demuxing virtual lanes and recovering lane markers on each virtual lane). Remux the virtual lanes bitwise into a serial stream to carry across the OTN, where the serial bit order is preserved. Demux at a bit level at the OTN egress into the appropriate number of LAN lanes depending on which PMD is chosen.
- Option B2 - Deskew the LAN virtual lanes at the OTN ingress (requires demuxing virtual lanes and recovering 64B/66B on each virtual lane). Remultiplex into a serial stream by assembling the 66B blocks in correct temporal order, resulting in a bitstream that looks like 10G Base-R but faster. At the OTN egress, demux the 66B blocks into 20 virtual lanes, remux bitwise into the required number of LAN lanes.

# Four lane 40 GbE interface





# 40 GbE PCS - What might 40 GbE Serial look like?



## Issues with using bit-muxed VLs for 40 GbE over OTN

---

- If the LAN is parallel, there is the same problem with skew budget as with considering the bit-muxed VL format for 100 GbE. (Note that for serial LAN, there is no issue with extra skew)
- There is also the problem that the bit-rate exceeds the capacity of standard ODU3.
- This would steer towards option B2 from the earlier slide, as transcoding proposals are based on a series of 66B blocks.

# Conclusions

- The concept of stealing from the IPG to make room for lane markers is not the same for the MLD(CTBI)/virtual lane architecture as it was for 10G Base-X architectures:
  - Lane marker size of 264 bits (4x66B) (40 GbE) or 1340 bits (20x66B) cannot be accommodated “in place” in the IPG the way a 40 bit lane marker (4x10B) can be inserted in a 10G Base-X interface
  - Lane markers interrupt a packet
  - Lane markers for 40 GbE and 100 GbE will redistribute IPG across the interface since a direct replacement of idles with lane markers and lane markers with idles is not possible
  - Since the LAN is not transparent with respect to IPG (not really different from 10G, but not widely recognized at 10G), some new definition is required as to what constitutes a transparent mapping over OTN
- The logical serial LAN format that would follow from the MLD architecture is a bit-mux of virtual lanes rather than a sequence of 66B blocks.
  - For parallel LAN, deskew is likely needed at the OTN ingress which requires recovery of 66B blocks
  - 66B blocks must also be recovered to perform transcoding for 40 GbE into ODU2
  - Option B2 appears to best meet the needs for a common method for OTN transport of parallel or serial 40 GbE and 100 GbE