40 GbE and 100 GbE PCS Considerations
November 2007 Update

Stephen J. Trowbridge
Alcatel-Lucent

Supporters

- Pete Anslow (Nortel)
- Frank Chang (Vitesse)
- Martin Carroll (Verizon)
- Ralf-Peter Braun (Deutsche Telekom)
- George Young (AT&T)
- Med Belhadj (Cortina)

PCS Considerations

- Additional information to trowbridge_02_0907.pdf 40 GbE and 100 GbE PCS considerations
- Ingress and Egress of OTN network should be free to use different Ethernet PMDs
- Need for a common understanding of the meaning of "transparent" transport of 40 GbE and 100 GbE
  - 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)
  - However, no serial PMDs are planned (initially) for 40 GbE or 100 GbE. So what does transparent service mean?

Key Questions

- Will lane alignment markers be transported over a serial LAN interface in the future?
- Should lane alignment markers be transported over a server layer such as OTN?
- Should lane alignment markers be inserted inband (steal bandwidth from the MAC by deleting from IPG) or out of band (increase the lane bitrate)?

PCS options based on whether lane alignment markers steal from the MAC bandwidth and whether or not they are carried across the OTN

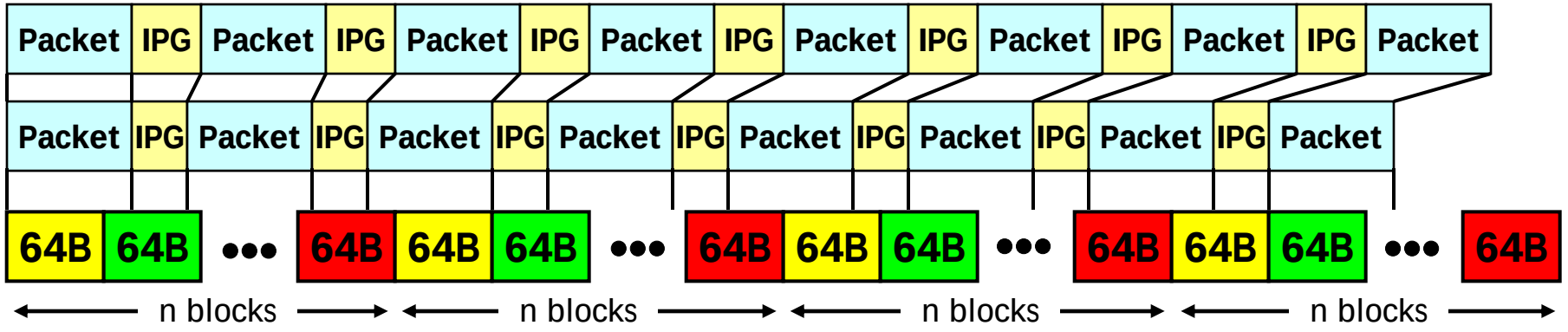
		OTN Interface Is space for lane alignment markers stolen from MAC bandwidth?			
		No Are lane alignment markers carried across OTN?		Yes Are lane alignment markers carried across OTN?	
		No	Yes	No	Yes
LAN interface Is space for lane alignment markers stolen from MAC bandwidth?	Yes	Option 1	Not likely	Option 2	Option 3
	No	Option 4	**Option 5 (new)**	Not likely	Not likely

Note: An implication of Options 2, 3, and 5 is that (virtual) lane alignment markers will ALWAYS be present, even when a serial PMD exists for 40 GbE or 100 GbE. This may be needed anyway, if only to perform the electrical deskew across the CTBI/CFBI.

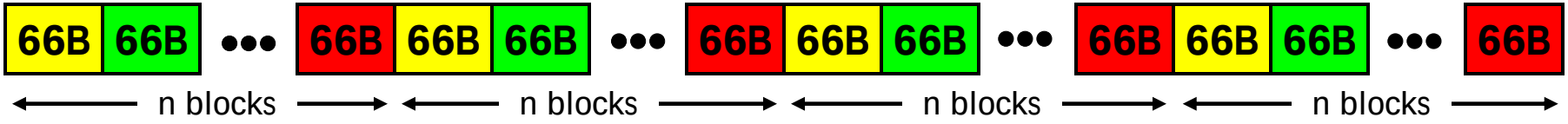
Lane alignment markers inserted by stealing from IPG

Options 1, 2, 3

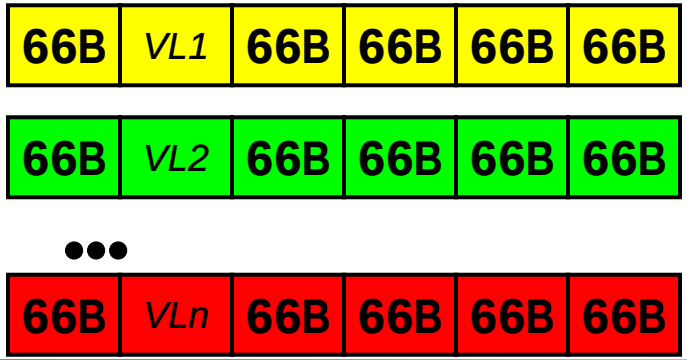
MII: Prior to PCS coding, reduce rate by 0.0061% by deleting from IPG



64B/66B encode:



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



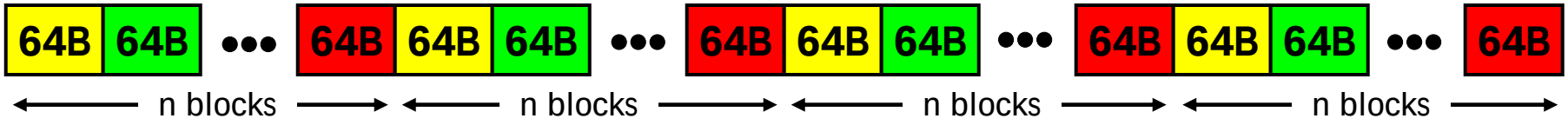
Lane rate = PCS encoded rate / n

Alcatel·Lucent

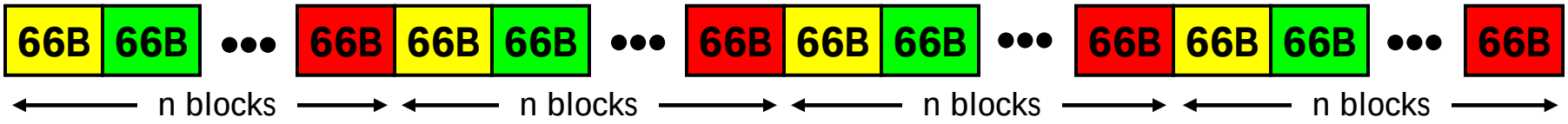
Lane alignment markers inserted by increasing lane rate

Options 4, 5

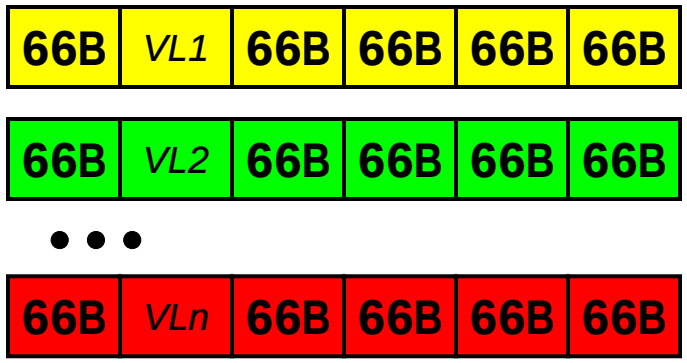
MII:



64B/66B encode:



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



Lane rate =
PCS encoded rate / n
x 1.000061

Option Summary - Where in the parallel LAN stack does the 64B/66B encoded data for OTN come from?

- 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
- 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
- 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**

See trowbridge_02_0907.pdf

Example Data flow for 100 GbE, 4 lane LAN interface with Options 4, 5 Increase lane rate to accommodate lane alignment markers

CGMII		100 Gb/s
Tx PCS	64B/66B coding	103.125 Gb/s
	Divide into 20 virtual lanes	20x5.15625 Gb/s
	Add lane alignment marker to each 16383 data or control 66B blocks per lane	20x5.15656473 Gb/s
Tx CTBI	Two virtual lanes per physical lane	10x10.31312946 Gb/s
Tx gearbox PMA Tx PMD Rx PMD Rx gearbox PMA	Five virtual lanes per physical lane	4x25.78282366 Gb/s
Rx CTBI	Two virtual lanes per physical lane	10x10.31312946 Gb/s
Rx PCS	Deskew virtual lanes, remove lane alignment markers	103.125 Gb/s
	Decode 64B/66B	100 Gb/s

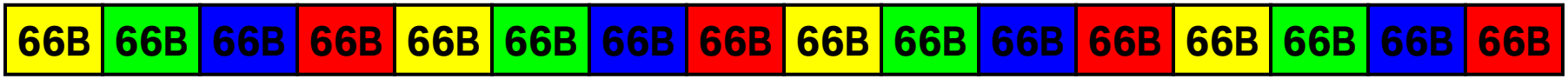
Data flow for 40 GbE, 4 lane LAN interface – Options 4, 5

Increase lane rate to accommodate lane alignment markers

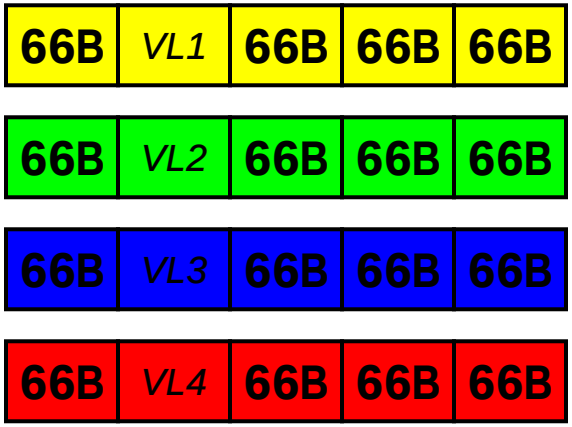
XLGMII		40 Gb/s
Tx PCS	64B/66B coding	41.25 Gb/s
	Divide into four virtual lanes	4x10.3125 Gb/s
	Add lane alignment marker to each 16383 data or control 66B blocks per virtual lane	4x10.3125 Gb/s
Tx XLFBI	One virtual lane per physical lane	4x10.31312946 Gb/s
Tx PMD Rx PMD	One virtual lane per physical lane	4x10.31312946 Gb/s
Rx XLFBI	One virtual lane per physical lane	4x10.31312946 Gb/s
Rx PCS	Deskew virtual lanes, remove lane alignment markers	41.25 Gb/s
	Decode 64B/66B	40 Gb/s

Options 3, 5 for 40 GbE – Deskew and keep lane alignment markers when mapping into OTN

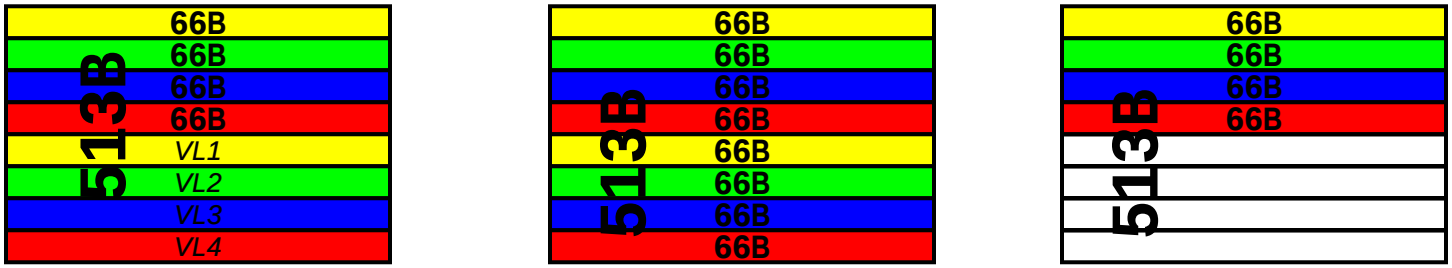
Delete from IPG, 64B/66B encode:



Inverse multiplex into lanes, add lane alignment markers:



Deskew, keep lane alignment markers, transcode to 512B/513B:



Option 3 - Lane alignment markers replace 0.0061% of IPG – same PCS encoded rate

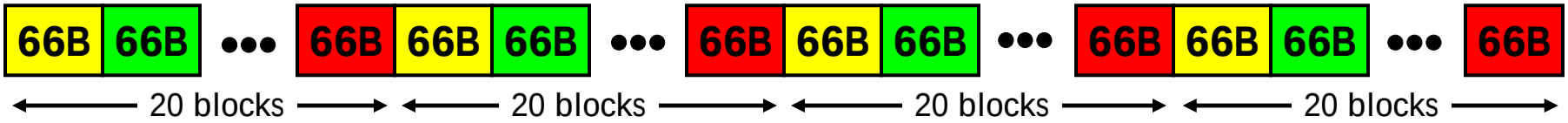
Option 5 – Lane rate increased by 0.0061% to accommodate lane alignment markers

Data flow for Adaptation of 40 GbE, 4 lane LAN interface into ODU3 – Option 5 – keep lane alignment markers

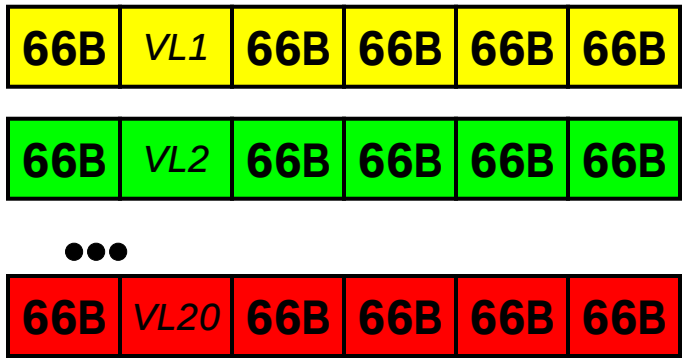
LAN PMD	One virtual lane per overrate physical lane with lane alignment markers	4x10.313129463 Gb/s
OPU3 adaptation	Deskew virtual lanes, keep lane alignment markers	41.252517854 Gb/s
	Transcode 64B/66B to 512B/513B	40.080571324 Gb/s
	Add Framing for 513B blocks	40.096197278 Gb/s
	Worst case +100ppm	40.100206897 Gb/s
OPU3 payload area	Worst case -20ppm	40.149716 Gb/s

Options 3, 5 for 100 GbE – Deskew, leaving lane alignment markers in place, and map into OTN

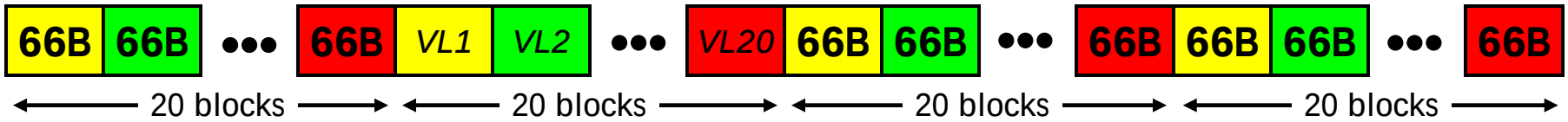
Delete from IPG, 64B/66B encode:



Inverse multiplex into 20 virtual lanes, add lane alignment markers:



Deskew, keep lane alignment markers, map into OPU4



Option 3 - Lane alignment markers replace 0.0061% of IPG – same PCS encoded rate

Option 5 – Lane rate increased by 0.0061% to accommodate lane alignment markers

Data flow for Adaptation of 100 GbE, 4 lane LAN interface into ODU3 – Option 5 – keep lane alignment markers

LAN PMD	Five virtual lanes per physical lane with lane alignment markers	4x25.782823659 Gb/s
OPU4 adaptation	Deskew virtual lanes, keep lane alignment markers	103.131294635 Gb/s
	Worst case +100ppm	103.141607764 Gb/s
OPU4	Least possible nominal bitrate given ±20ppm clock tolerance	103.143670638 Gb/s

Conclusions

- Several viable options to establish relationship between the multi-lane LAN PCS and the 64B/66B or 512B/513B encoded bitstream to be mapped into OTN
- The following options were viewed most favorably at the ITU-T Q11/15 meeting in Shenzhen due to the ability to provide a transparent service:
  - Option 4: deskew and remove lane alignment markers which are accommodated by increasing the LAN lane rate rather than stealing from the MAC
  - Option 5: deskew but keep the virtual lane alignment markers in the OTN transported bitstream
- These options require that virtual lane marking will be done even when IEEE 802.3 defines a serial PMD for 40 GbE or 100 GbE;
  - but this may be necessary anyway for electrical deskew across a CTBI or XLFBI type interface
- Care should be taken in choosing the alignment word format to avoid expanding the 66B codeword space or to create difficulty for transcoding to 512B/513B in 40 GbE