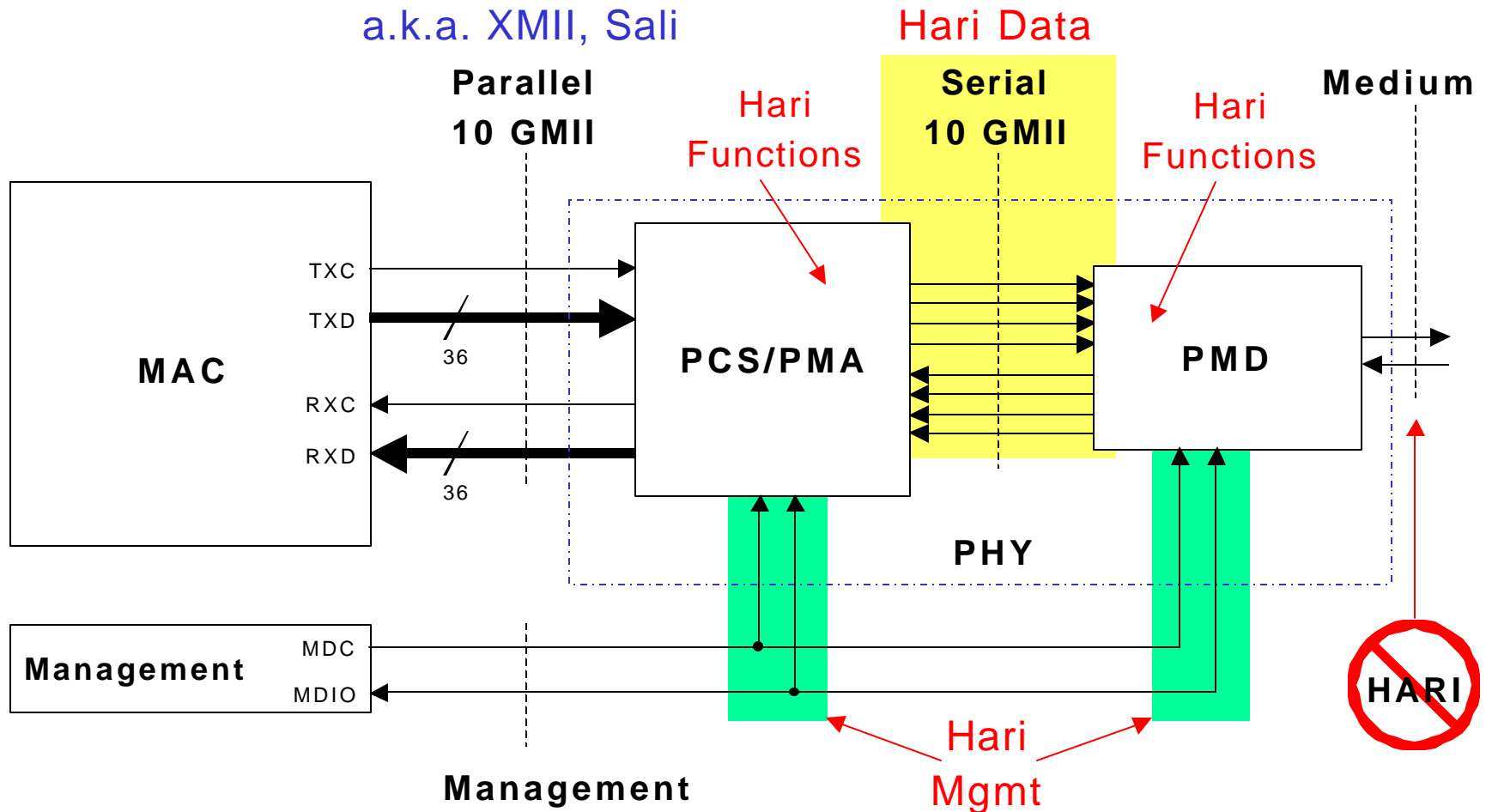


IEEE 802.3ae Task Force

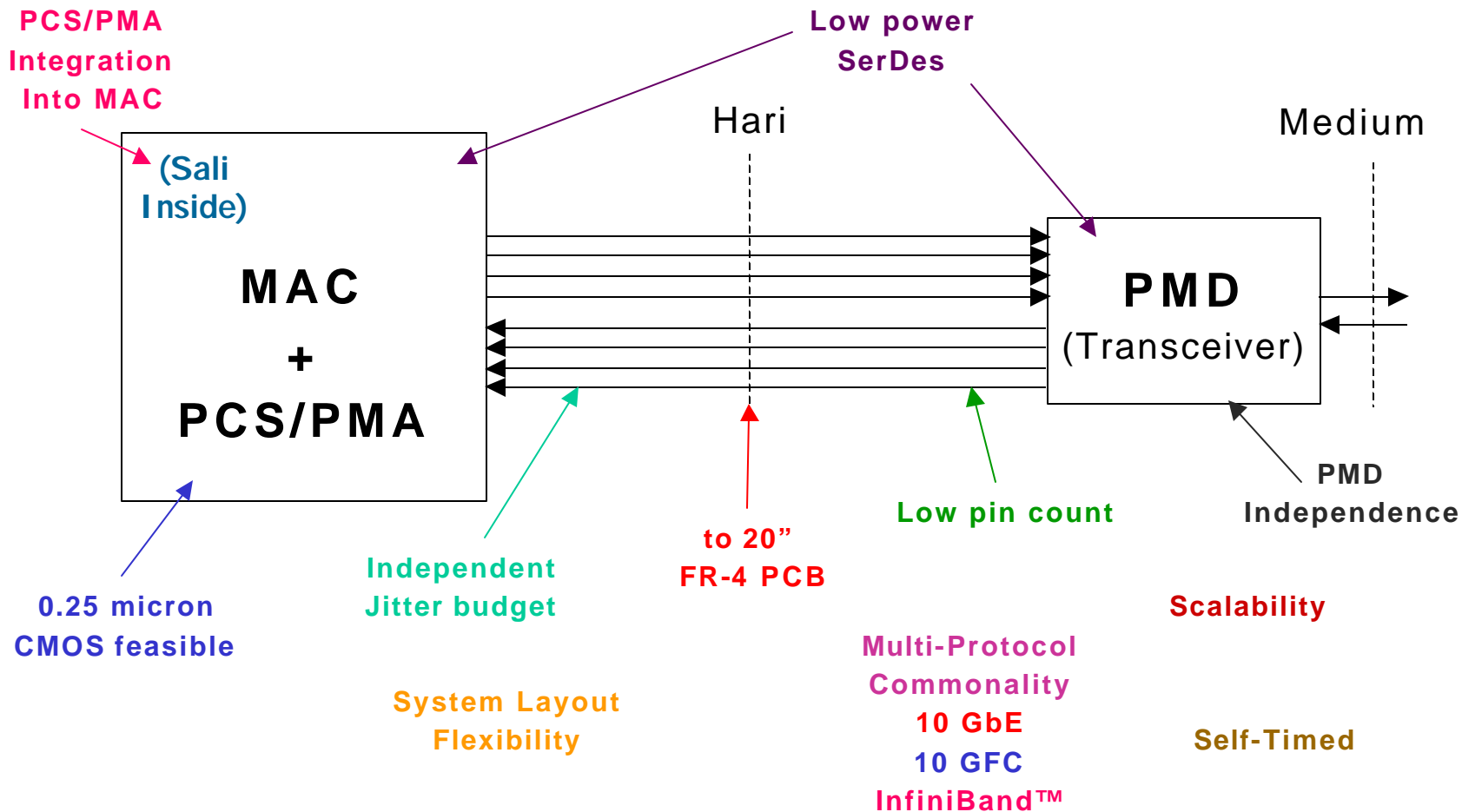
Hari Coding Issues & Proposal

Dallas, TX
January 18, 2000
Richard Taborek Sr.
nSerial Corporation
408-845-6102
rtaborek@earthlink.net

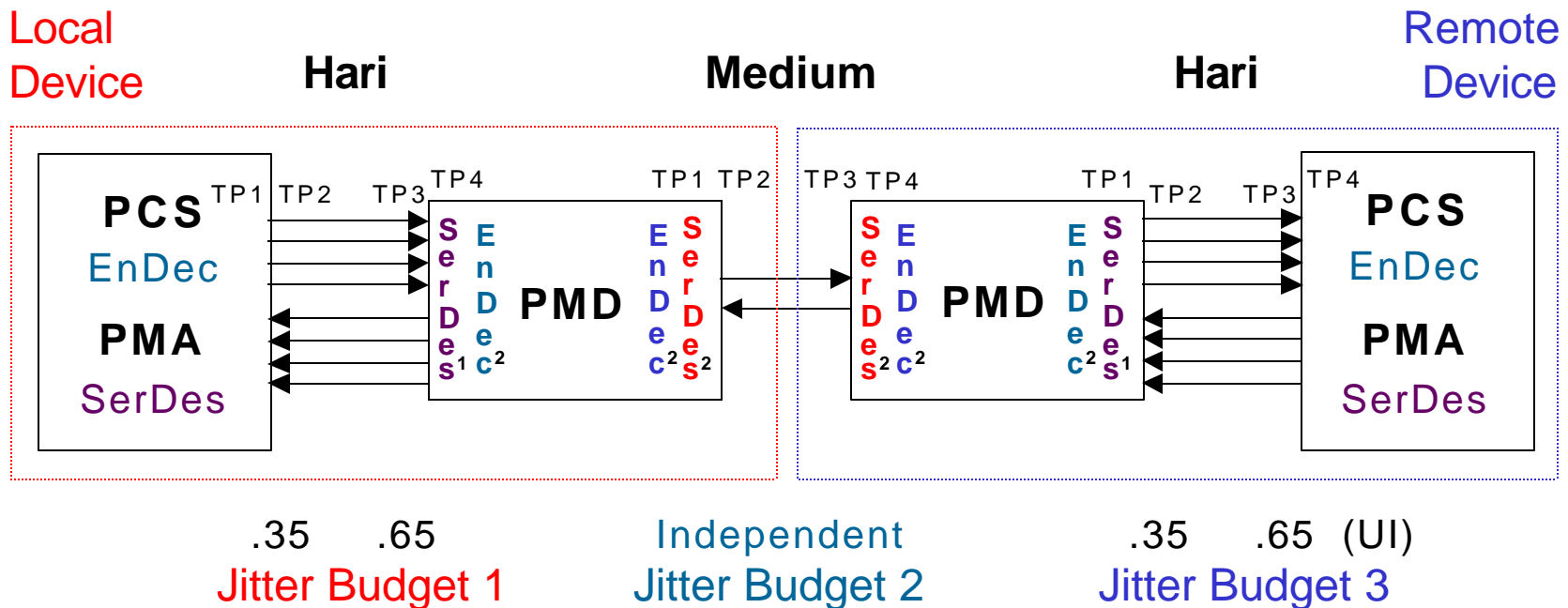
Where does Hari fit in?



Hari Benefits



10G Link Architecture/Jitter Budget

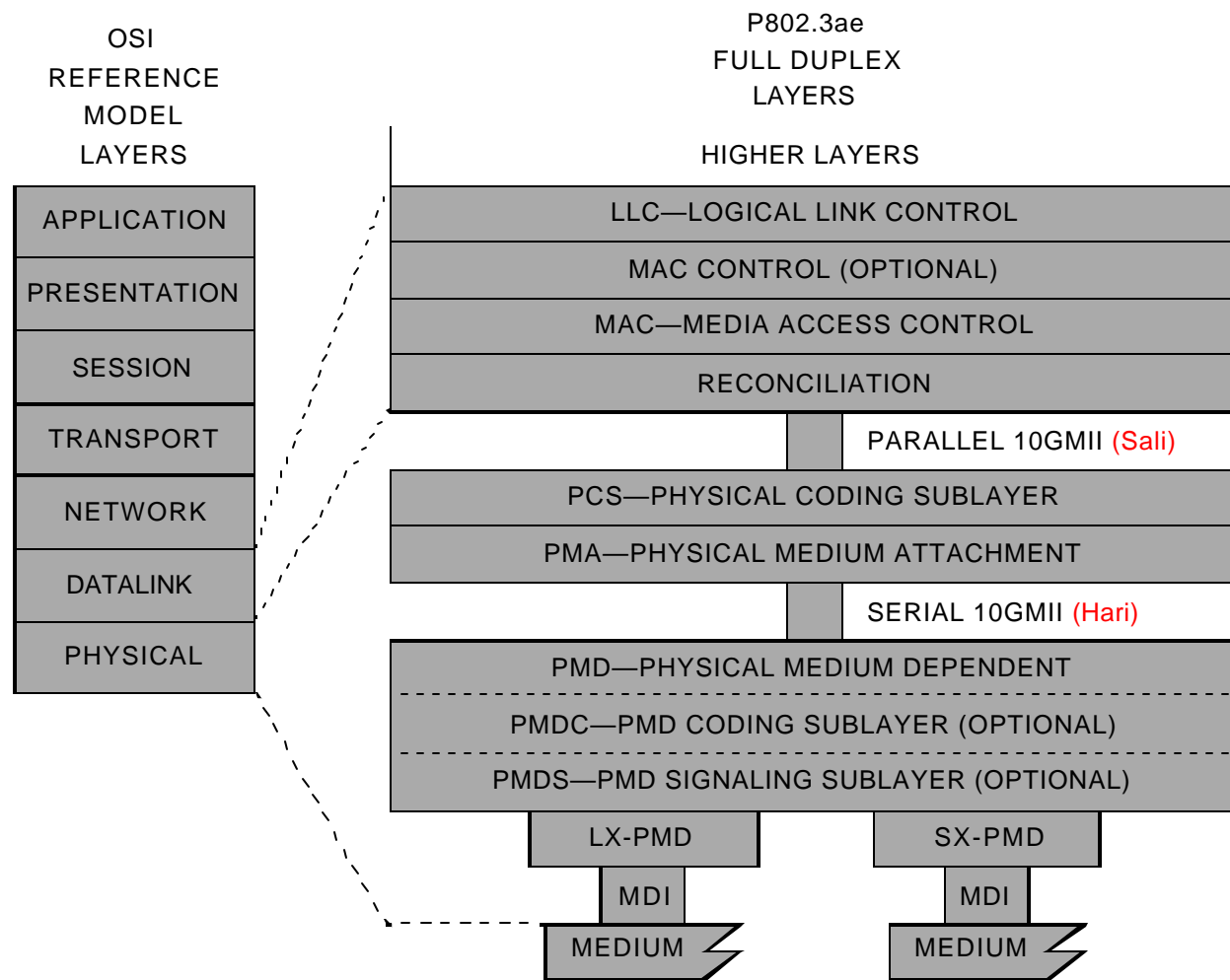


- Medium Jitter Budget is independent of Hari
 - ◆ Requires Hari functionality in PMD
- Hari-based PMD simplifies 10 GbE link development

¹ Optional Hari retiming functionality

² Optional PMD recoding/signaling

OSI Model/10 GbE Proposed Layers



Hari Coding Functions & Features

- Mapping: Direct to/from Parallel 10 GMII (Sali)
 - ◆ PMD direct to/from 64B/66B or PAM5x4; Same as WWDM
- Link Synchronization: Comma, /K/, /K28.5/
 - ◆ PCS generates /K/ columns during IPG in /K/R/ pattern
- Clock Tolerance Compensation: Skip, /R/, /K28.0/
 - ◆ PCS generates /R/ columns during IPG in /K/R/... pattern
 - ◆ Receiver inserts/removes /R/ columns
 - ◆ Multiple Clock Domains are optional, but may help control jitter
- Lane Deskew - Alignment of Received Columns
 - ◆ Required for Skip and Low-Latency data processing
 - ◆ Align to /K/ columns present in PCS Idle pattern

Direct Sali-to-Hari Mapping

Sali (a.k.a. XMII, Parallel 10 GMII)

D<0:7>	I	I	S	d _p	d	d	---	d	d	d	d _f	I	I	I	I	I
D<8:15>	I	I	d _p	d _p	d	d	---	d	d	d _f	T	I	I	I	I	I
D<16:23>	I	I	d _p	d _p	d	d	---	d	d	d _f	I	I	I	I	I	I
D<24:31>	I	I	d _p	d _s	d	d	---	d	d	d _f	I	I	I	I	I	I

Hari (a.k.a. Serial 10 GMII)

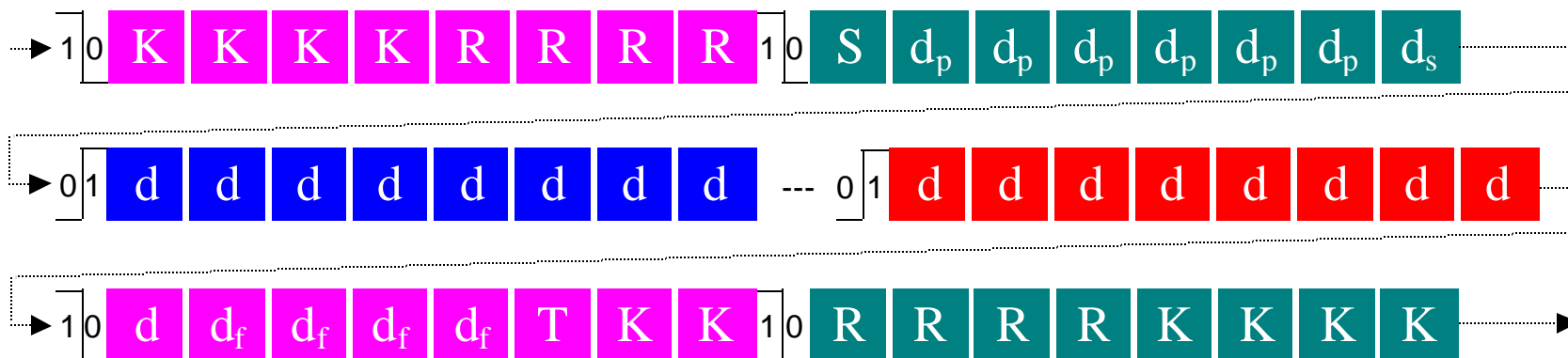
Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K	R
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K	R
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K	R
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K	R

Direct Hari-to-64B/66B Mapping

Hari columns partitioned into 64B/66B sub-frames

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K	R
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K	R
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K	R
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K	R

64B/66B sub-frames in serial transmission order



64B/66B code-group representations shown. Specific sub-frame mapping describe in walker_1_0100.pdf

Direct Hari-to-PAM5×4 Mapping

Sali (a.k.a. XMII, Parallel 10 GMII)

D<0:7>	I	I	S	d _p	d	d	---	d	d	d	d _f	I	I	I	I	I
D<8:15>	I	I	d _p	d _p	d	d	---	d	d	d _f	T	I	I	I	I	I
D<16:23>	I	I	d _p	d _p	d	d	---	d	d	d _f	I	I	I	I	I	I
D<24:31>	I	I	d _p	d _s	d	d	---	d	d	d _f	I	I	I	I	I	I

Hari (a.k.a. Serial 10 GMII)

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K	R
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K	R
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K	R
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K	R

PAM5×4

R	R	R	S	d _p	d _p	---	d	d	d	---	d _f	T	K	K	R	R	R	R
---	---	---	---	----------------	----------------	-----	---	---	---	-----	----------------	---	---	---	---	---	---	---

PAM5x4 code-group representations shown. Specific symbol mapping describe in taborek_2_1199.pdf

10 GbE Hari Idle Encoding

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K

- **/K/R/** repeating Idle pattern
 - ◆ **/K/ = /K28.5/**, contains comma
 - Used for Link Synchronization, Lane Deskew, EOP padding
 - ◆ **/R/ = /K28.0/**, disparity neutral
 - **/R/** columns (Skips) may be inserted/removed to compensate for clock tolerance differences
- “Even/Odd” column alignment is superfluous

10 GbE Hari SOP Encoding

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K

- /S/ Start of Packet Delimiter
 - ◆ /S/ = /K27.7/
 - ◆ Also serves as Lane 0 ID
 - Useful for addressing Serial PMD “lane rotation”

10 GbE Hari Data Encoding

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K

- **/d/** Packet Data
 - ◆ **/d_p/** Preamble; **/d_s/** Preamble SFD; **/d_f/** FCS
- Supports WAN PHY Preamble replacement, but...
 - ◆ ...Only if Column-Striping is used
 - ◆ Word-Striping requires commas in Preamble

10 GbE Hari EOP Encoding

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K

- /T/ End of Packet (EOP) Delimiter
 - ◆ /T/ = /K29.7/
 - ◆ EOP padded with /K/

10 GbE Hari Error Encoding

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K
Lane 2	K	R	d _p	d _p	d	d	---	d	E	d _f	K	R	K	R	K
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K

- /E/ The “dreaded” Error code-group
 - ◆ Same as GbE /V/ Void code-group
 - ◆ /E/ = /K30.7/
 - ◆ Signaled when an error is detected in the received signal or needs to be forced into the transmitted signal

10 GbE - PMD Inserts /R/ column

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	K	R	K
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	K	R	K
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	K	R	K
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	K	R	K

PMD Inserts /R/ column here 

Lane 0	K	R	S	d _p	d	d	---	d	d	d	d _f	R	R	K	R	K
Lane 1	K	R	d _p	d _p	d	d	---	d	d	d _f	T	R	R	K	R	K
Lane 2	K	R	d _p	d _p	d	d	---	d	d	d _f	K	R	R	K	R	K
Lane 3	K	R	d _p	d _s	d	d	---	d	d	d _f	K	R	R	K	R	K

Hari Coding Issues (here we go!)

1. To B 8B/10B or not to B
2. Column-Striping or Word-Striping
3. EMI concerns with bandwidth constrained codes
4. Multi-Protocol support
5. Multi-PMD support
6. Initialization pattern for Deskew/Link Sync

1. Hari Transmission Code

- Other codes have been proposed for Hari
 - ◆ MB810, 1000BASE-T derivatives, etc.
- 8B/10B is ubiquitous, robust, simple
- Well known by Ethernet equipment designers
- Low gate-count/complexity Encode/Decode
- 10 GbE mappings close to 1000BASE-X PCS
- No other code proven to be significantly better

∴ No reason to change. Stick with 8B/10B.

2. Column-Striping or Word-Striping

- Hari WILL work with either striping method
 - ◆ Economic, NOT Technical feasibility, IS the issue
- 10 Gbps PMD/Protocol independent Hari is Key
 - ◆ Column-Striping works best for all: 10 GbE, 10 GFC, InfiniBand™, OIF, Serial LAN PMD @ 10.3125 Gbaud, WAN PHY, MAS PMD, WWDM PMD
- Low Power discrete and core SerDes are needed
 - ◆ Only enabled by **new** SerDes designs
 - PCS/PMA co-location enables SerDes simplification
 - Move traditional high-power SerDes functions to parallel logic

∴ **Go for the \$Green\$. Choose Column-Striping**

3. EMI concerns with 8B/10B

- 8B/10B is a bandwidth constrained code
- EMI concerns abound with GbE at 1.25 Gbps
 - ◆ Hari 10 GbE EMI concerns are 10X GbE
 - ◆ This should say: *Houston... we have a problem...*
- This Engineering problem can we well contained
 - ◆ Hari architecture, Column-Striping are containment keys
 - ◆ IEEE 1394b has already addressed this problem
 - Solution directly applicable to 8B/10B encoded links.

∴ We're great at leveraging technology: P1394b

4. Multi-Protocol support

- ~10 Gbps PHYs include 10 GbE, 10 GFC, InfiniBand™, SONET OC-192c
 - Most 10 Gbps high-volume PHYs should be similar
 - ◆ LAN, SAN, MAN/WAN access
 - The IEEE 802.3ae Task Force is setting direction
 - ... For cost effective 10 Gbps PHYs
 - ◆ We need to make the proper technology choices
 - ◆ Make them in light of other industry activities/directions
- ∴ Support Hari, 8B/10B, Column-Striping for all**

5. Multi-PMD support

- Hari works for all PMDs, including WAN
 - Hari is a strategic chip-to-chip interconnect
 - Hari is NOT a PMD
 - ◆ Hari supports the ATTACHMENT of the MAC to PMD
 - Hari simplifies 10 Gbps link designs
 - ◆ Greatly increases medium jitter budget
 - ◆ Enhanced integration, scalability, flexibility
 - ◆ With today's chip technology, Hari is no more complex than yesterday's parity bit, but buys a lot more!
- ∴ All PMDs benefit from Hari. Standardize it!**

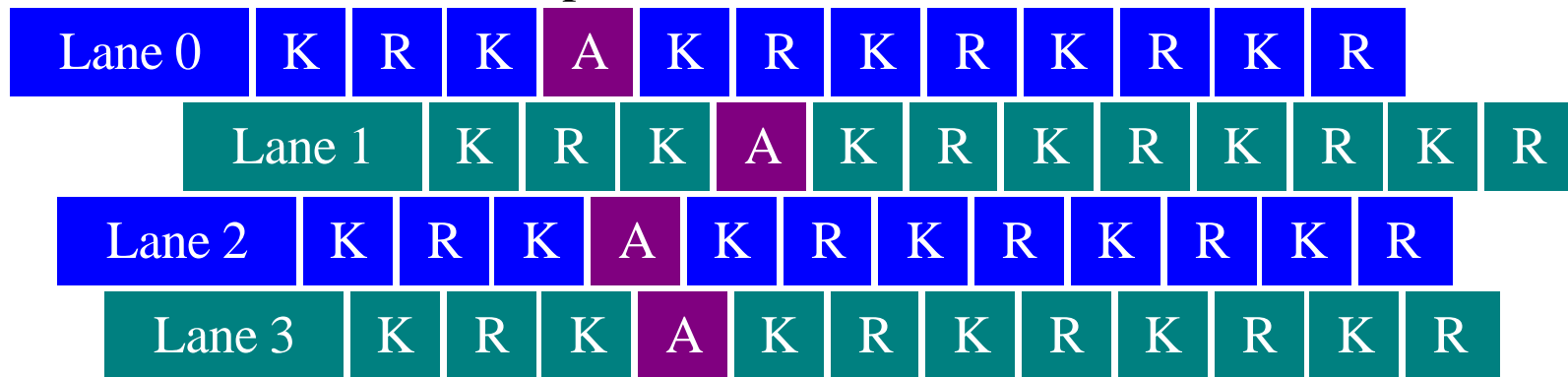
6. Deskew/Link Sync Initialization

- Two problems with Deskew and Initialization
 - a) DeSer process may increase effective skew
 - See dedrick_1_0100.pdf
 - b) KR Idle is only 40-bits long during Initialization
- A simple robust Idle solves both problems
- Solution proposed in dedrick_1_0100.pdf, additionally
 - ◆ Define Align code-group, /A/ = /K28.3/
 - ◆ /A/ Columns replace /R/ Columns as 1st full IPG column and every 16th IPG column, if applicable. Not deletable.
 - ◆ /K/ code-groups always pad out column containing EOP
 - ◆ Receiver aligns to /A/K/, /A/ provides comma “cursor”

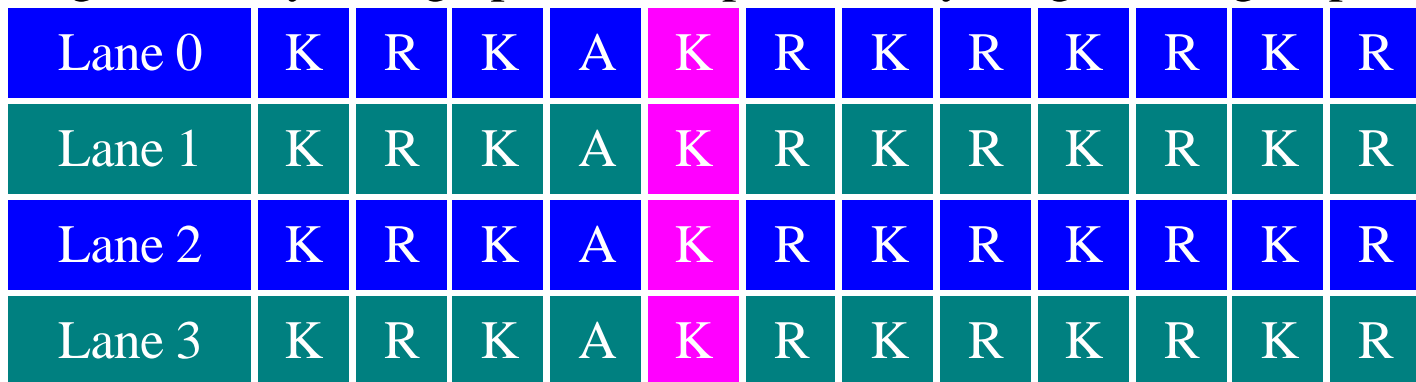
:: Add Alignment Function to Hari

New 10 GbE Receiver Deskew

Skewed data at receiver input. Skew ~18 bits



Align lanes by lining up commas prefixed by Align code-groups



Hari Coding Summary

- Simple rules, Protocol/Application independence
 - ◆ Common coding rules for 10 GbE, IB, FC, WAN PHY
- Cost-effective “system interface” for all PMDs
- Leverages high-reliability, ubiquitous 8B/10B code
 - ◆ Low gate-count/complexity Encode/Decode, well known
- Virtually identical to 1000BASE-X PCS
- Column aligned data enables simplest Rx process
 - ◆ Enables lowest power SerDes design
 - ◆ No serialization delays, low speed clocking
 - ◆ Implementation flexibility (e.g. 4/8/16 octets/cycle)
- Overall best match for simple 10G System Interface