Impact of MLD Path Delay Proposals on Timestamps

Denny Wong, Xilinx

IEEE P802.3cx Teleconference

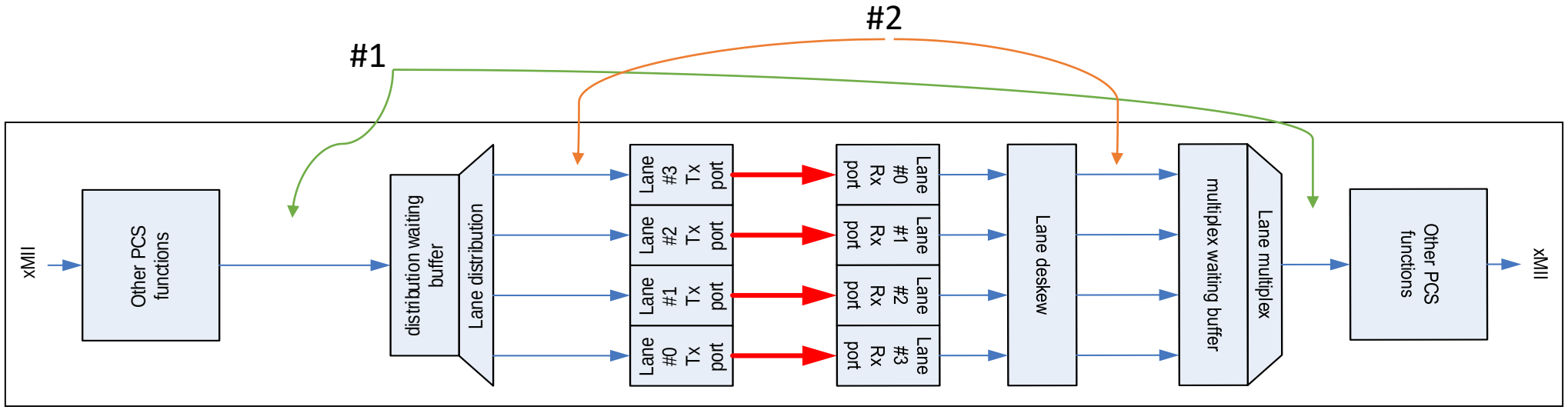
October 20, 2020

Introduction

- Different approaches have been discussed regarding the handling of path delay variance due to multi-PCS lane distribution. Aside from the mechanics, the impact on the actual timestamps must be considered.

Proposed Solutions

- Two main approaches were proposed and discussed in tse_3cx_02_520 and he_3x_01_0920 (and others)
  - Approach 1: Option C + Method 2, where Tx and Rx use a constant delay instead of the actual delay, because the sum of the end-to-end intrinsic delay is constant.
  - Approach 2: Option B + Method 1, where Tx and Rx use time at MDI (or accurately compensate TS delay from xMII to MDI).



Approach 1

Block Time	640	ps	time of one block's worth of data (i.e. 64 bits of uncoded data)
Tx Option	C		Blocks on NxPCS lane Tx MDI are aligned, xMII to MDI delay is different for every Tx PCS lane, blocks on Tx PCS lanes have different timestamps because a predetermined constant lane distribution delay is used for all lanes
Rx Option	2		Each Rx lane's MDI to xMII delay is different, but is accounted for as a predetermined constant delay because it can be balanced by a mirror delay on the Tx side
Link delay	0	ps	assumes the same delay for all lanes
Number of lanes	20		number of PCS lanes (max = 20)
Resulting time err	**0**	ps	

Lane		actual arvl time		actual distrib dly		actual dept time		actual arrival time		actual deskew dly		actual merge dly		actual arrival time		
	0		0		12160		12160		12160		0		0		12160	
	1		640		11520		12160		12160		0		640		12800	
	2		1280		10880		12160		12160		0		1280		13440	
	3		1920		10240		12160		12160		0		1920		14080	
	4		2560		9600		12160		12160		0		2560		14720	
	5		3200		8960		12160		12160		0		3200		15360	
	6		3840		8320		12160		12160		0		3840		16000	
	7		4480		7680		12160		12160		0		4480		16640	
	8		5120		7040		12160		12160		0		5120		17280	
	9		5760		6400		12160		12160		0		5760		17920	
	10		6400		5760		12160		12160		0		6400		18560	
	11		7040		5120		12160		12160		0		7040		19200	
	12		7680		4480		12160		12160		0		7680		19840	
	13		8320		3840		12160		12160		0		8320		20480	
	14		8960		3200		12160		12160		0		8960		21120	
	15		9600		2560		12160		12160		0		9600		21760	
	16		10240		1920		12160		12160		0		10240		22400	
	17		10880		1280		12160		12160		0		10880		23040	
	18		11520		640		12160		12160		0		11520		23680	
	19		12160		0		12160		12160		0		12160		24320	

Lane				modelled PHY dly		dept timestamp					modelled PHY dly		arrival timestamp		measured link dly	
	0				6080		6080					6080		6080		0
	1				6080		6720					6080		6720		0
	2				6080		7360					6080		7360		0
	3				6080		8000					6080		8000		0
	4				6080		8640					6080		8640		0
	5				6080		9280					6080		9280		0
	6				6080		9920					6080		9920		0
	7				6080		10560					6080		10560		0
	8				6080		11200					6080		11200		0
	9				6080		11840					6080		11840		0
	10				6080		12480					6080		12480		0
	11				6080		13120					6080		13120		0
	12				6080		13760					6080		13760		0
	13				6080		14400					6080		14400		0
	14				6080		15040					6080		15040		0
	15				6080		15680					6080		15680		0
	16				6080		16320					6080		16320		0
	17				6080		16960					6080		16960		0
	18				6080		17600					6080		17600		0
	19				6080		18240					6080		18240		0

All lanes have unique timestamps

Approach 2

Block Time	640	ps	time of one block's worth of data (i.e. 64 bits of uncoded data)
Tx Option	B		Blocks on NxPCS lane Tx MDI are aligned, xMII to MDI delay is different for every Tx PCS lane, blocks on every Tx PCS lane have same timestamp because per-lane distribution delay is accounted for
Rx Option	1		Each Rx lane's MDI to xMII delay is different and is accounted for as such
Link delay	0	ps	assumes the same delay for all lanes
Number of lanes	20		number of PCS lanes (max = 20)
Resulting time err	**0**	ps	

Lane	actual arvl time	actual distrib dly	actual dept time	actual arrival time	actual deskew dly	actual merge dly	actual arrival time
0	0	12160	12160	12160	0	0	12160
1	640	11520	12160	12160	0	640	12800
2	1280	10880	12160	12160	0	1280	13440
3	1920	10240	12160	12160	0	1920	14080
4	2560	9600	12160	12160	0	2560	14720
5	3200	8960	12160	12160	0	3200	15360
6	3840	8320	12160	12160	0	3840	16000
7	4480	7680	12160	12160	0	4480	16640
8	5120	7040	12160	12160	0	5120	17280
9	5760	6400	12160	12160	0	5760	17920
10	6400	5760	12160	12160	0	6400	18560
11	7040	5120	12160	12160	0	7040	19200
12	7680	4480	12160	12160	0	7680	19840
13	8320	3840	12160	12160	0	8320	20480
14	8960	3200	12160	12160	0	8960	21120
15	9600	2560	12160	12160	0	9600	21760
16	10240	1920	12160	12160	0	10240	22400
17	10880	1280	12160	12160	0	10880	23040
18	11520	640	12160	12160	0	11520	23680
19	12160	0	12160	12160	0	12160	24320

Lane	modelled PHY dly	dept timestamp	modelled PHY dly	arrival timestamp	measured link dly
0	12160	12160	0	12160	0
1	11520	12160	640	12160	0
2	10880	12160	1280	12160	0
3	10240	12160	1920	12160	0
4	9600	12160	2560	12160	0
5	8960	12160	3200	12160	0
6	8320	12160	3840	12160	0
7	7680	12160	4480	12160	0
8	7040	12160	5120	12160	0
9	6400	12160	5760	12160	0
10	5760	12160	6400	12160	0
11	5120	12160	7040	12160	0
12	4480	12160	7680	12160	0
13	3840	12160	8320	12160	0
14	3200	12160	8960	12160	0
15	2560	12160	9600	12160	0
16	1920	12160	10240	12160	0
17	1280	12160	10880	12160	0
18	640	12160	11520	12160	0
19	0	12160	12160	12160	0

All lanes have same timestamp

Impact on Timestamps

- Because “Approach 1” uses a constant PHY delay for each PCS block, it has a timestamp granularity of one PCS block
- “Approach 2” uses actual PHY delay, which means all PCS lanes at MDI are parallel. This results in a timestamp granularity of (number of lanes x PCS block time), since all lanes will have the same timestamp.
  - Unique timestamps can be inferred if lane number is known, but this requires the information be provided.

Granularity at various rates

Rate	PCS Block Time	Lanes	Granularity Approach 1	Granularity Approach 2
25GE	2560ps	1	2.56ns	2.56ns
50GE	1280ps	4	1.28s	5.12ns
100GE	640ps	20	0.64ns	12.80ns
200GE	320ps	8	0.32ps	2.56ns
400GE	160ps	16	0.16ps	2.56ns

Summary

- “Approach 1” achieves the highest timestamp granularity, and is more suitable for high precision applications (Class C or better). Class D (5ns max TE) is impossible to reach for some rates with “Approach 2”.
- Even if different methods are used on the Tx and Rx sides, the error introduced is still less than the granularity of “Approach 2” (assuming the side using “Approach 1” uses half of the constant MLD delay on each of the Tx and Rx sides).
  - This should be considered when deciding where to assign the MLD delay (between Rx and Tx) in “Approach 1”.

Thank You!