MPSoC'16

July 12, 2016, Nara, Japan

Haar-based Interconnect Coding for Energy Effective, Reliable Data Transport

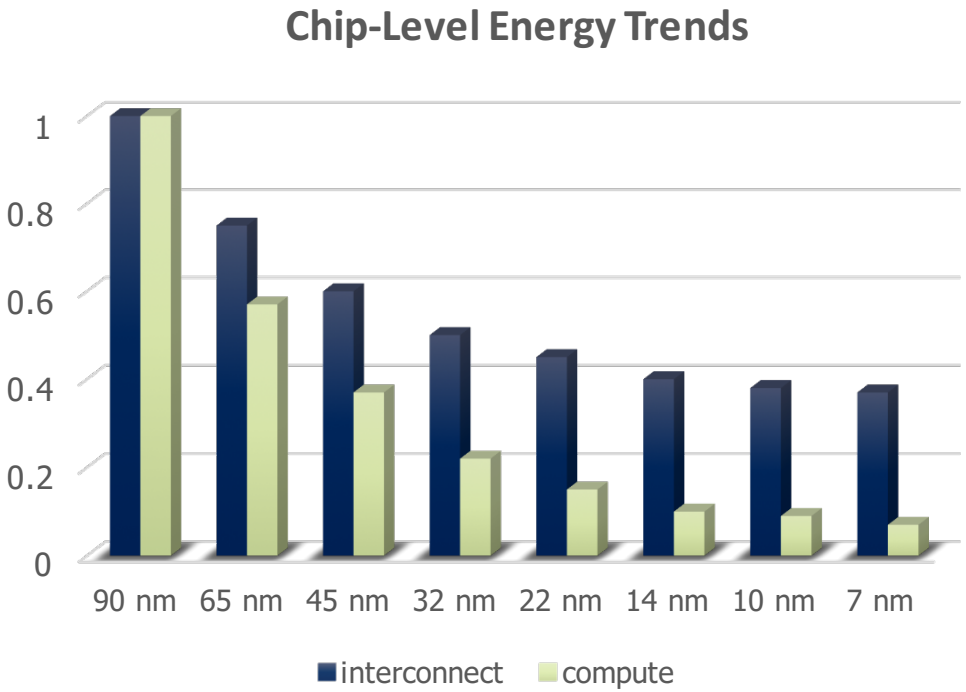
Sorin D. Cotofana

Computer Engineering Laboratory
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology (TU Delft), The Netherlands

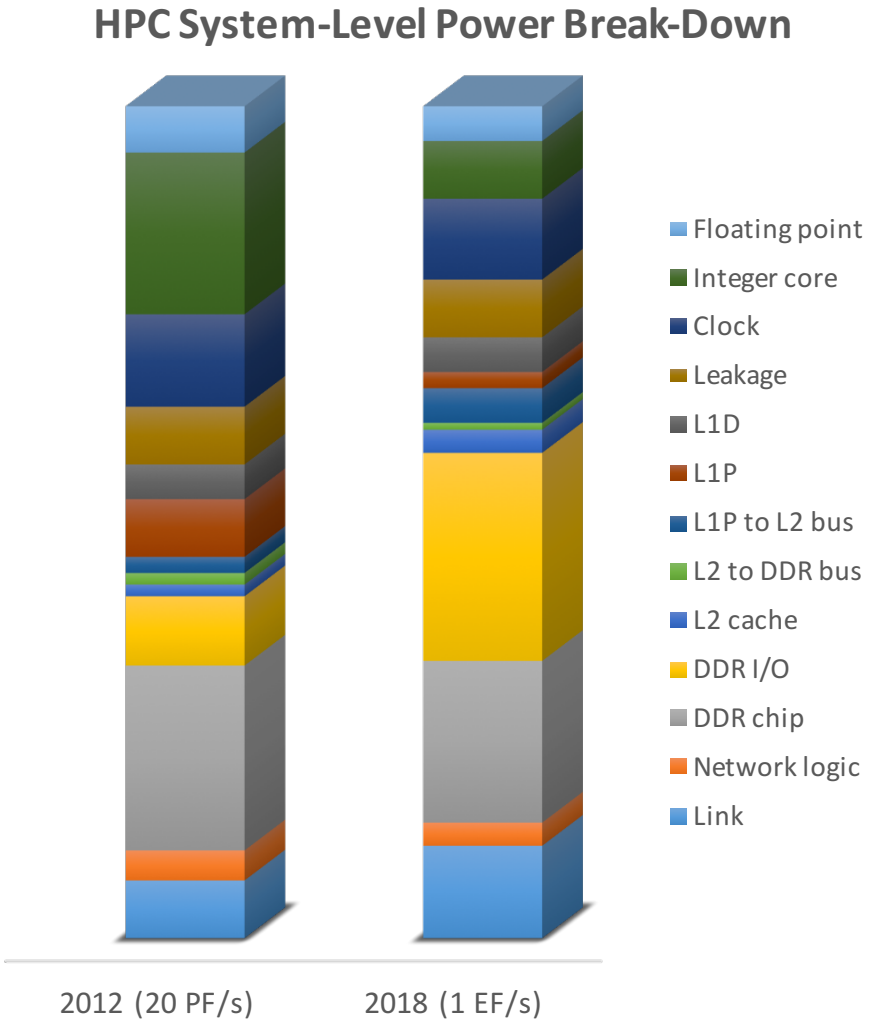
Overview

- The Interconnect Challenge
- Haar Codec
- Error Detection and Correction
- Evaluation
- Conclusions

Interconnect Escalating Challenges



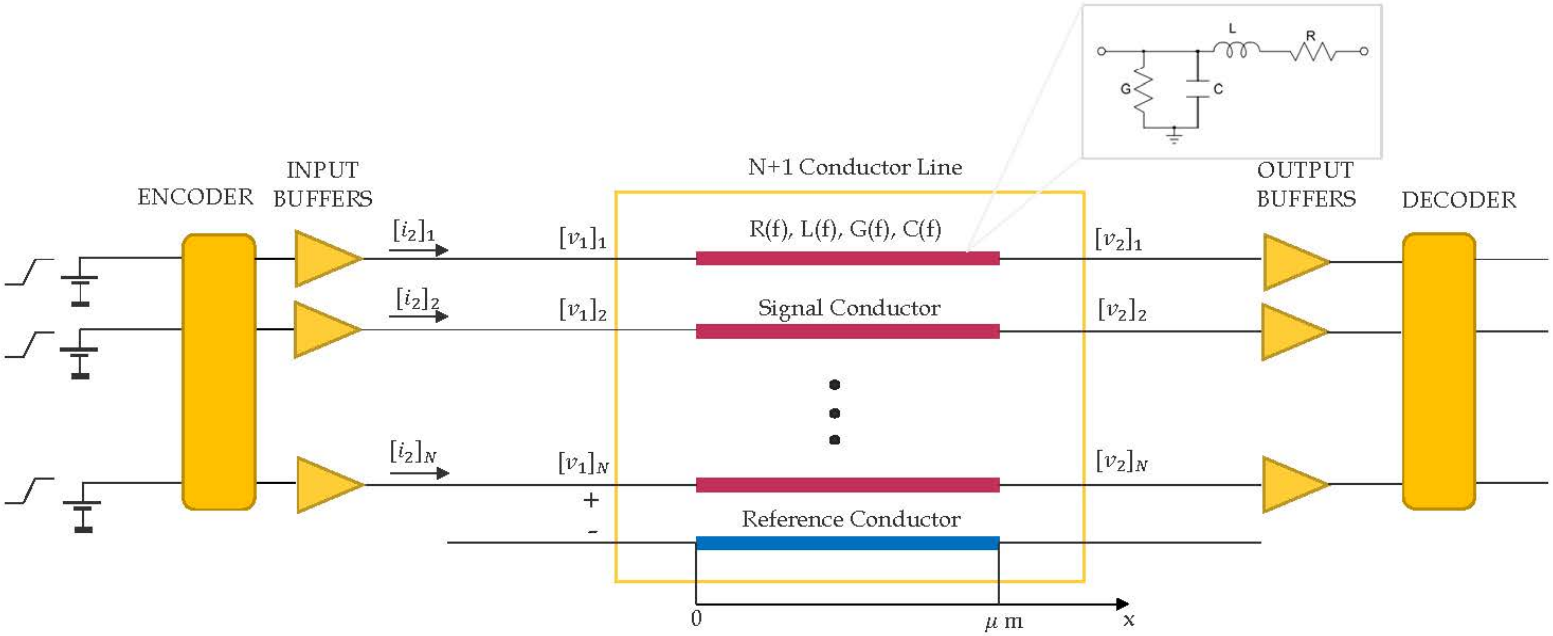
Source: Borkar, S., "Exascale Computing – A Fact or A Fiction?", IPDPS, 2013.



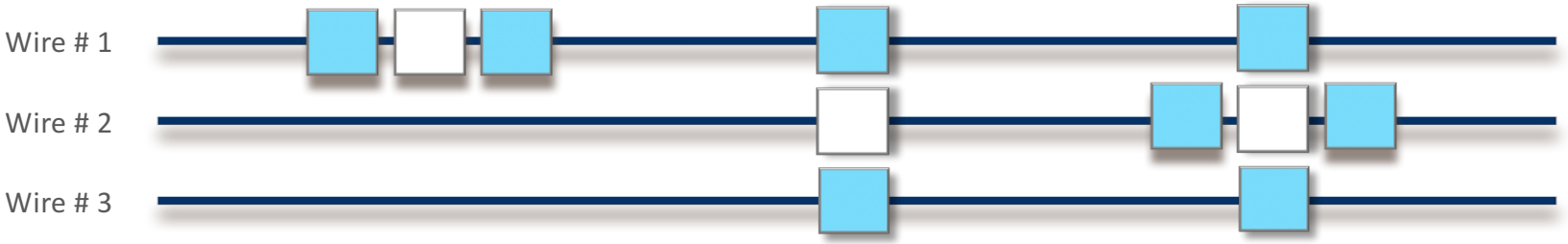
Source: Nair, R., "Active Memory Cube", 2nd Workshop on Near Data Processing, 2014.

- Communication dominates power
- Increased variability is penalizing performance and reliability

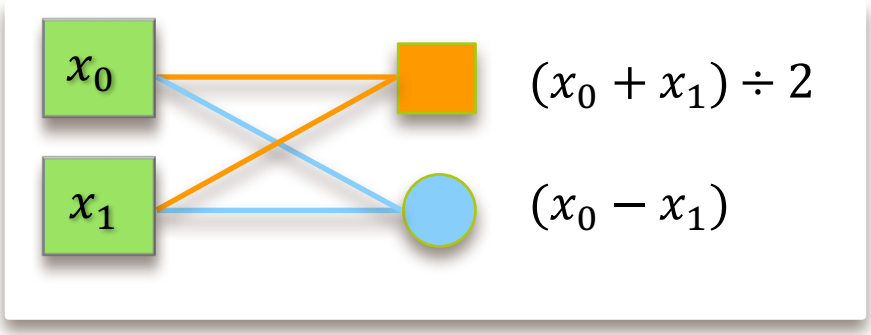
Interconnect Coding



Undesirable Patterns:



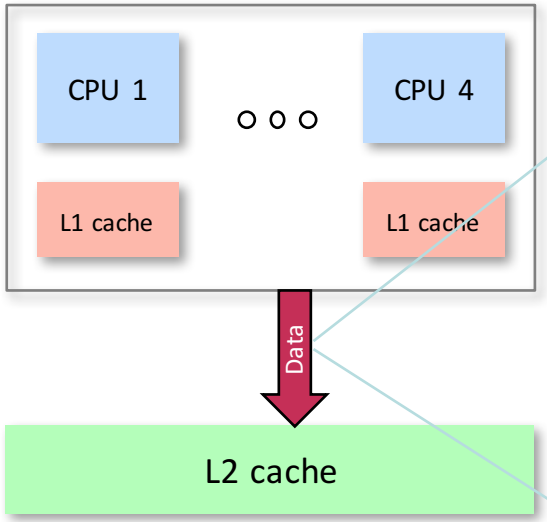
Haar Codec

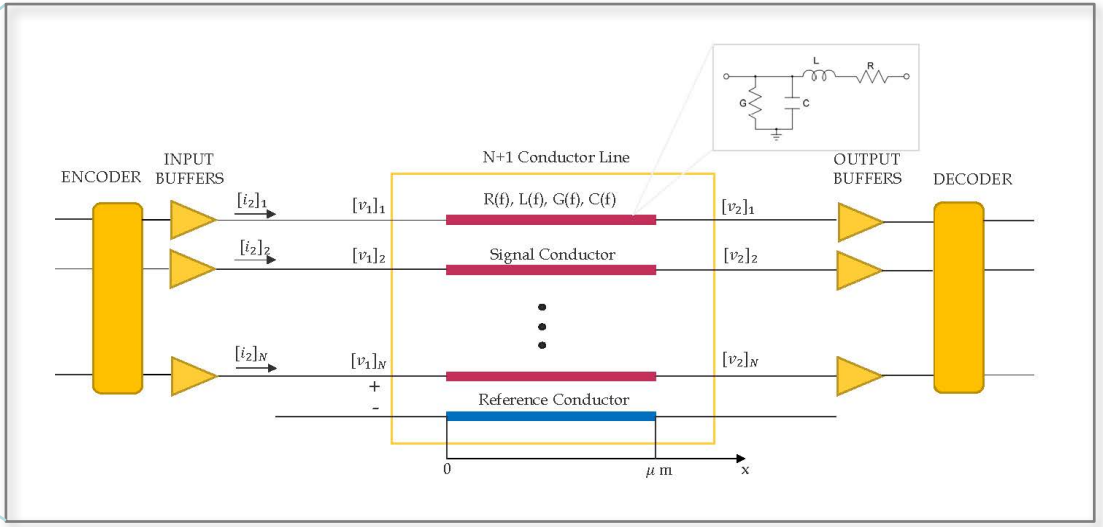


$$S^{(0)} = x_0 \wedge x_1$$
$$C_1^{(0)} = x_1$$
$$C_1^{(0)} = x_0 \vee x_1$$

$$\hat{x}_0 = S^{(0)} \oplus C_1^{(0)} \oplus C_0^{(0)}$$
$$\hat{x}_1 = C_1^{(0)}$$

SoC w/ Haar Assisted Interconnect





Frequency (GHz)	2
Issue & commit width	8
Int/FP instruction queue	32
Reorder buffer size	40
Load/store queue entries	16
L1 size (kB)/associativity	32/8
L1 hit latency (cycles)	2
L2 size (MB)/associativity	4/8
L2 hit latency CMOS (cycles)	12
Cache line size (bytes)	64

- SPEC CPU2000: 164.gzip; 176.gcc; 181.mcf
- 10.000 byte workload per benchmark

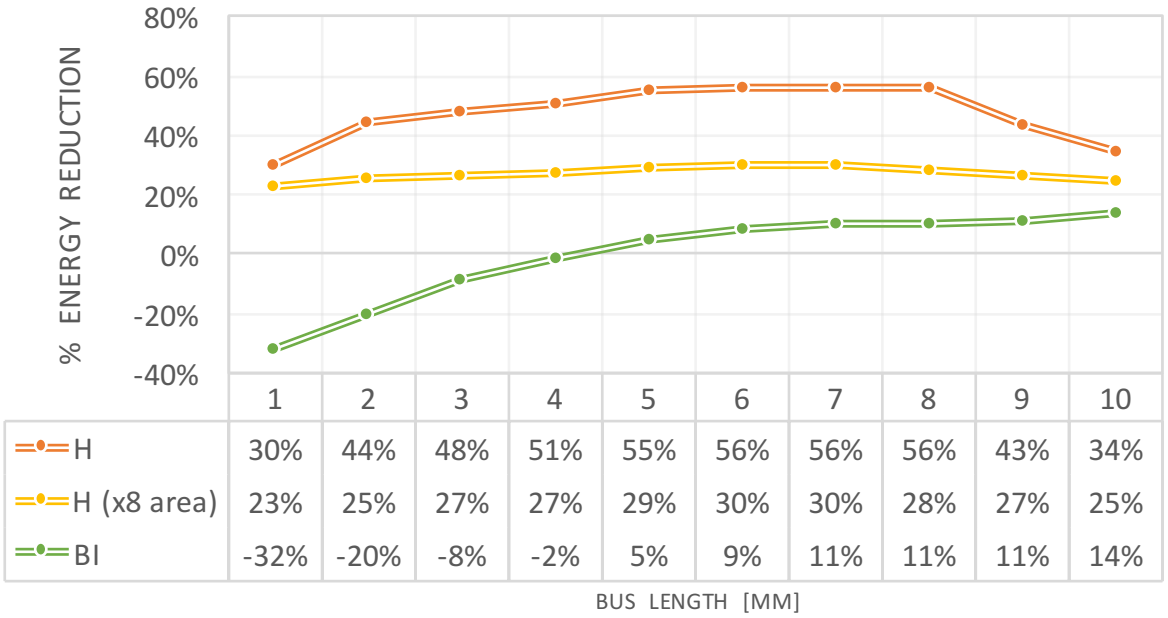
40nm interconnect
Bus length [mm]: 1, 2, …, 10
Bus width: 8, 16, …, 512

Statistics

- Data Arrival Profile
- Max. Frequency
- Energy
- Area

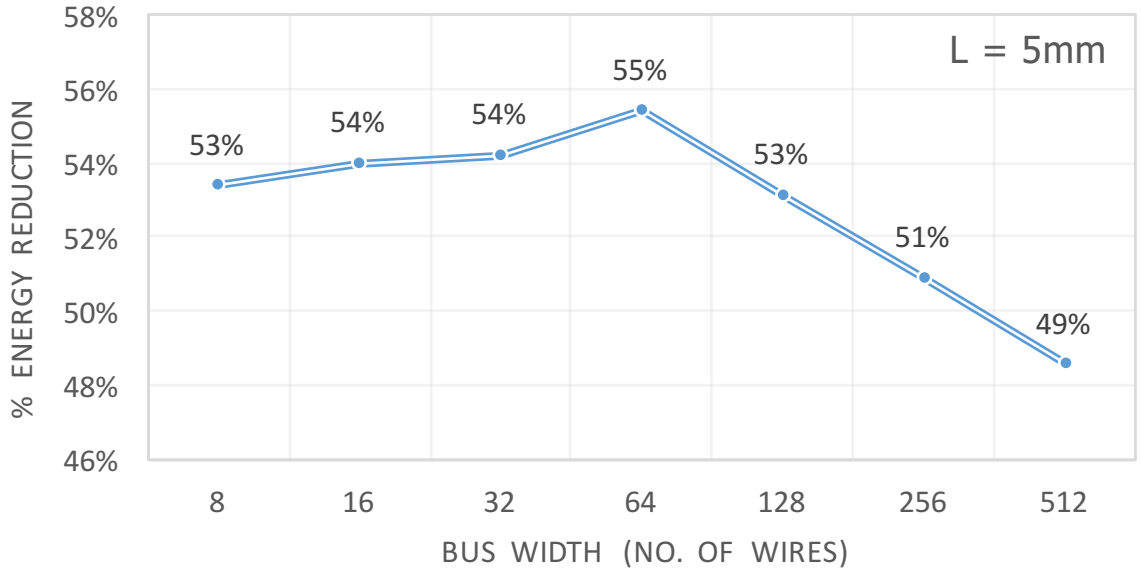
Transmission Configuration

	Wire Number											
	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12
Ref	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7				
H; H(x8 area)	$S^{(0)}$	$C_1^{(0)}$	$C_0^{(0)}$	$S^{(1)}$	$C_1^{(1)}$	$C_0^{(1)}$	$S^{(2)}$	$C_1^{(2)}$	$C_0^{(2)}$	$S^{(3)}$	$C_1^{(3)}$	$C_0^{(3)}$

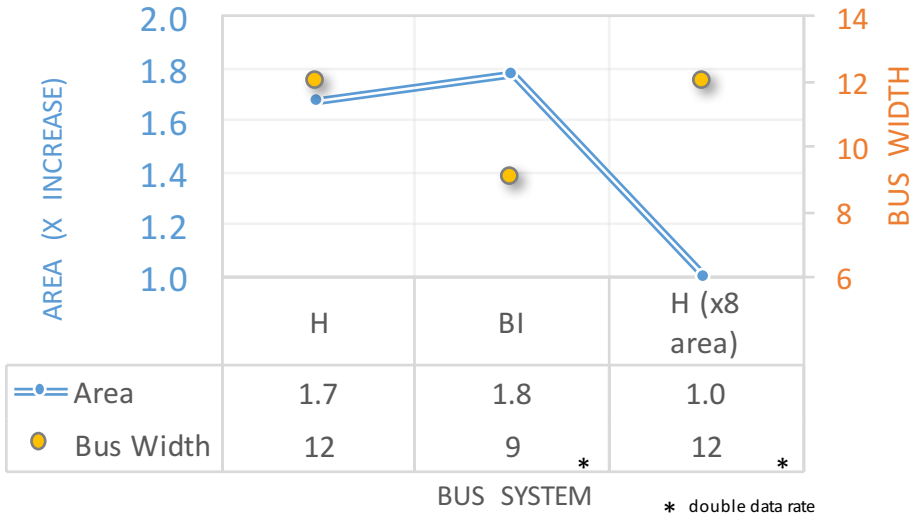


	1	2	3	4	5	6	7	8	9	10
H	30%	44%	48%	51%	55%	56%	56%	56%	43%	34%
H (x8 area)	23%	25%	27%	27%	29%	30%	30%	28%	27%	25%
BI	-32%	-20%	-8%	-2%	5%	9%	11%	11%	11%	14%

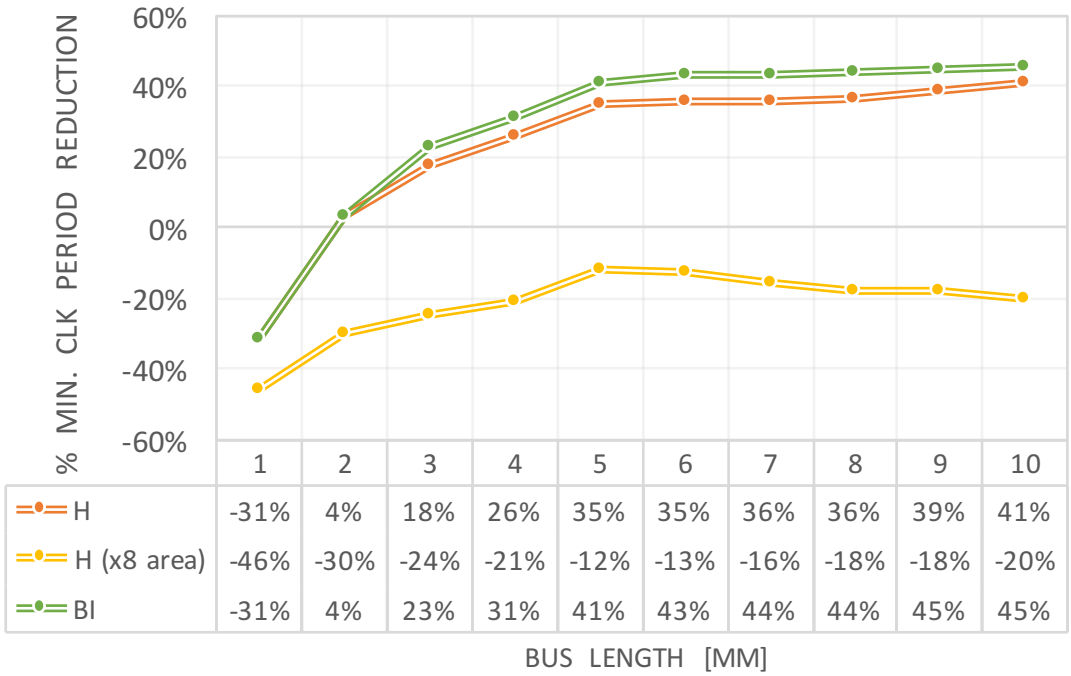
Energy vs. Bus Length



Energy vs. Bus Width



Min. CLK Period



Area Footprint

Error Detection and Correction

x_0	x_1	TX end (m_0)	RX end ($m_0 \oplus \varepsilon$)	$\hat{x}_0$	$\hat{x}_1$
0	0	(0 0 0)	(1 0 0)	1 (F)	1 (F)
			(0 1 0)		1 (F)
			(0 0 1)		0
0	1	(0 1 1)	(1 1 1)	1 (F)	0 (F)
			(0 0 1)		0 (F)
			(0 1 0)		1
1	0	(0 0 1)	(0 0 0)	0 (F)	0
			(0 1 1)		1 (F)
			(1 0 1)		1 (F)
1	1	(1 0 0)	(1 1 0)	0 (F)	1
			(1 0 1)		0 (F)
			(0 0 0)		0 (F)

To protect $\{x_0, x_2, x_4, x_6\}$:

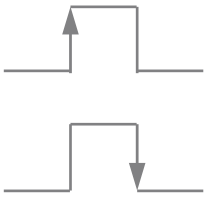
$$E_1 = x_0 \oplus x_2 \oplus x_6 \qquad E_3 = x_2 \oplus x_4 \oplus x_6$$
$$E_2 = x_0 \oplus x_2 \oplus x_4 \qquad E_4 = x_0 \oplus x_4 \oplus x_6$$

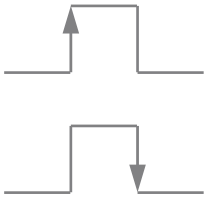
To recover $\{x_1, x_3, x_5, x_7\}$:

$$E_5 = C_1^{(0)} \oplus C_1^{(1)} \oplus C_1^{(2)} \oplus C_1^{(3)}$$

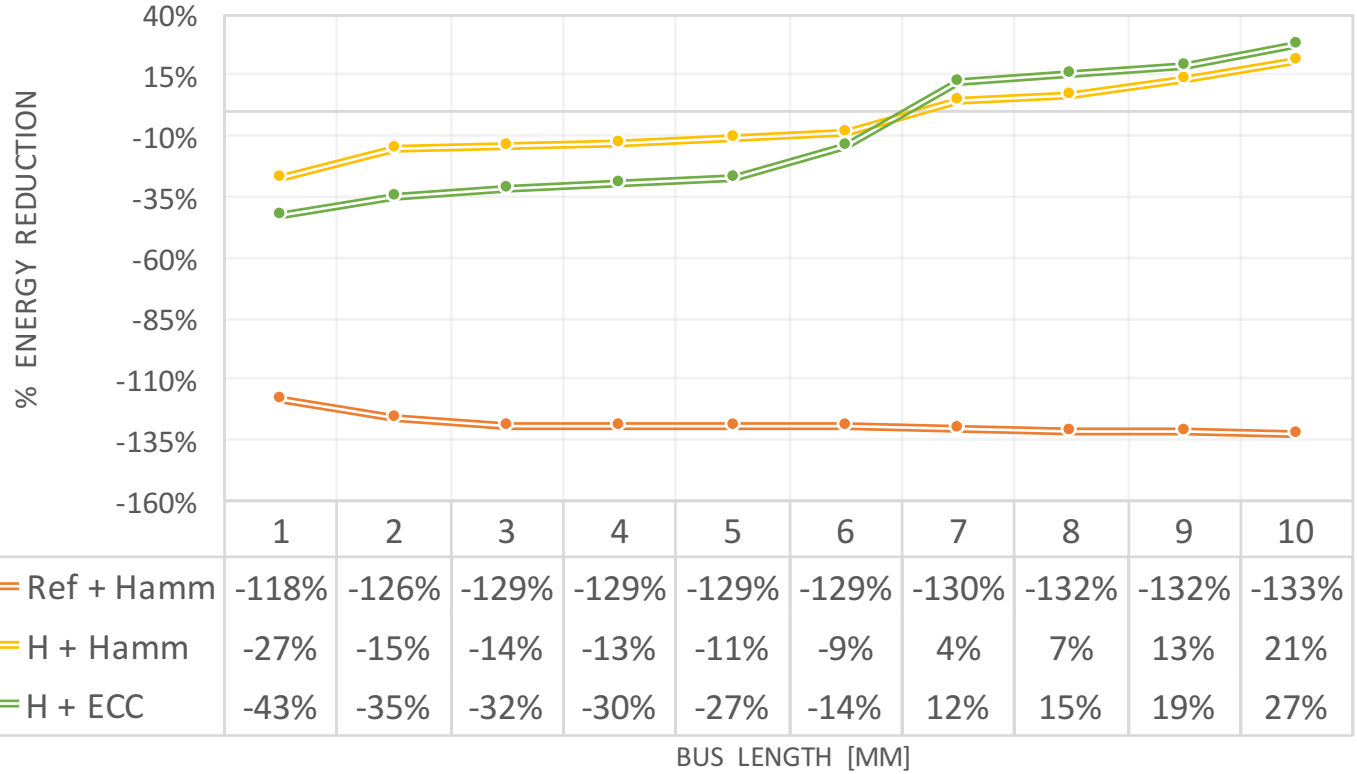
Transmission Configuration

	Wire Number											
	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12
Ref	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7				
Ref + Hamm	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	H_1	H_2	H_2	H_3
H + Hamm	$S^{(0)}$	$C_1^{(0)}$	$C_0^{(0)}$	$S^{(1)}$	$C_1^{(1)}$	$C_0^{(1)}$	H_1	H_2				
	$S^{(2)}$	$C_1^{(2)}$	$C_0^{(2)}$	$S^{(3)}$	$C_1^{(3)}$	$C_0^{(3)}$	H_3	H_4				
H + ECC	$S^{(0)}$	$C_1^{(0)}$	$C_0^{(0)}$	$S^{(1)}$	$C_1^{(1)}$	$C_0^{(1)}$	E_1	E_2	E_5			
	$S^{(2)}$	$C_1^{(2)}$	$C_0^{(2)}$	$S^{(3)}$	$C_1^{(3)}$	$C_0^{(3)}$	E_3	E_4				

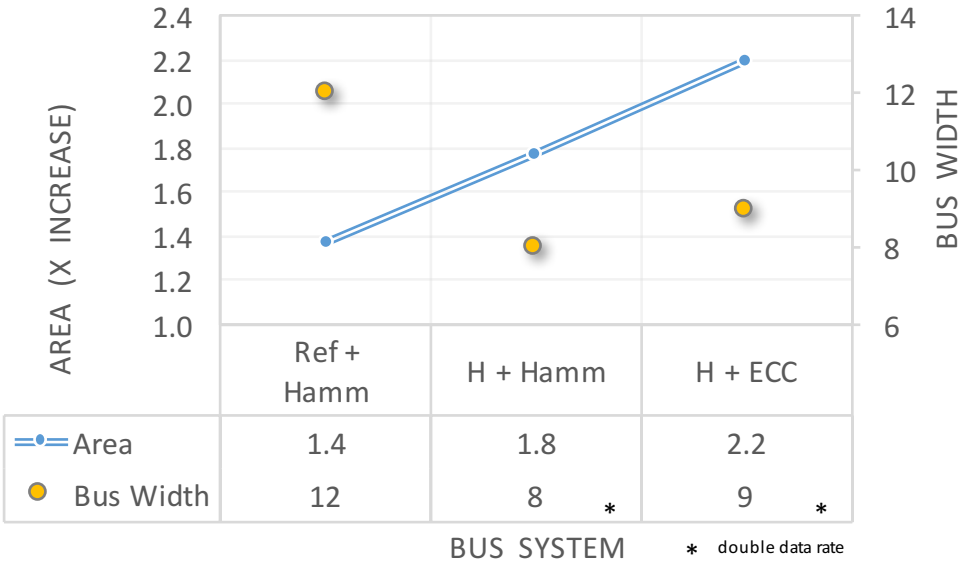




Energy

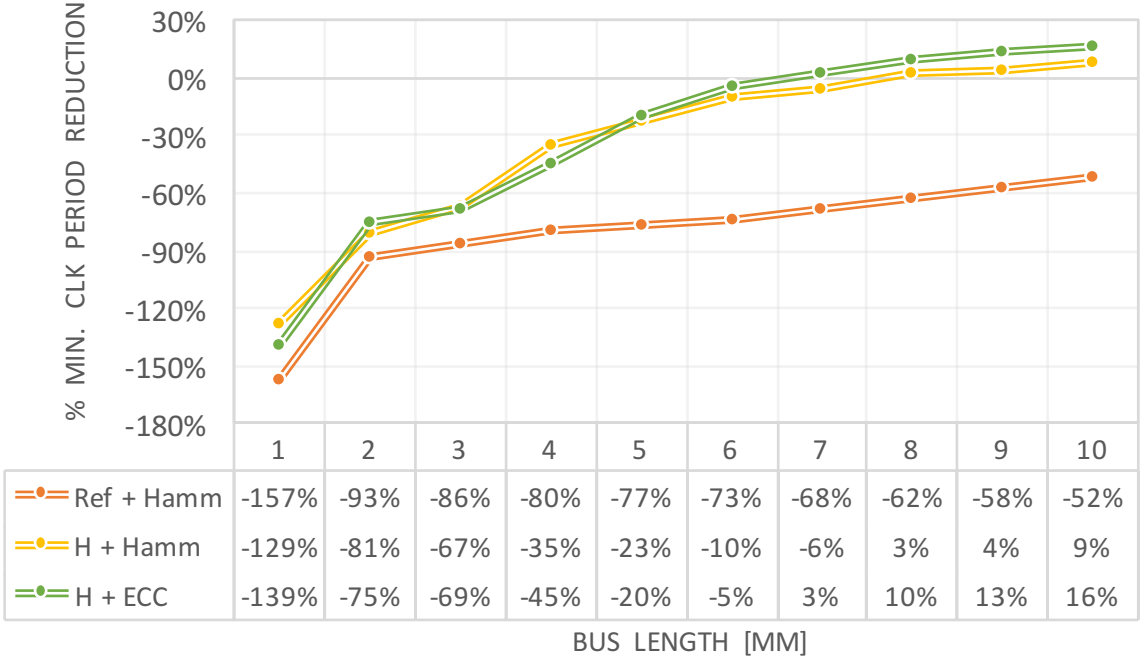


	1	2	3	4	5	6	7	8	9	10
Ref + Hamm	-118%	-126%	-129%	-129%	-129%	-129%	-130%	-132%	-132%	-133%
H + Hamm	-27%	-15%	-14%	-13%	-11%	-9%	4%	7%	13%	21%
H + ECC	-43%	-35%	-32%	-30%	-27%	-14%	12%	15%	19%	27%



Area Footprint

Min. CLK Period



Conclusions

- ECC & Haar Codec Augmented Interconnects:
  - Energy Effective
  - Faster
  - Fault Tolerant (SECDED)
- Synergistic Multi-Level Communication Architecture Change of Paradigm Concern


Thank you