

ESE532: System-on-a-Chip Architecture

Day 27: December 4, 2019
Defect Tolerance



Today

- Reliability Challenges
- Defect Tolerance
 - Memories
 - Interconnect
 - FPGA
- FPGA Variation and Energy
 - (time permitting)

Message

- At small feature sizes, not viable to demand perfect fabrication of billions of transistors on a chip
- Modern ICs are like snowflakes
 - Everyone is different, changes over time
- Reconfiguration allows repair
 - Finer grain → higher defect rates
 - Tolerate variation → lower energy

Intel Xeon Phi Offerings

Part #	Cores
7290F	72
7250F	68
7230F	64

<http://www.intel.com/content/www/us/en/products/processors/xeon-phi/xeon-phi-processors.html>

Intel Xeon Phi Offerings

Part #	Cores
7290F	72
7250F	68
7230F	64

Is Intel producing 3 separate chips?

Preclass 1 and Intel Xeon Phi Offerings

Part #	Cores
7290F	72
7250F	68
7230F	64

Cost ratio between 72 and 64 processor
assuming fixed mm² per core?

Intel Xeon Phi Pricing

CHOOSE YOUR OPTIMIZATION POINT

	CORES	GHZ	MEMORY	FABRIC	DDR4	POWER ²	RECOMMENDED CUSTOMER PRICING ²
7290 Best Performance/Node	72	1.5	16GB 7.2 GT/s	Yes	384GB 2400 Mhz	245W	\$6254
7250 Best Performance/Watt	68	1.4	16GB 7.2 GT/s	Yes	384GB 2400 Mhz	215W	\$4876
7230 Best Memory Bandwidth/Core	64	1.3	16GB 7.2 GT/s	Yes	384GB 2400 Mhz	215W	\$3710
7210 Best Value ¹	64	1.3	16GB 6.4 GT/s	Yes	384GB 2133 Mhz	215W	\$2438

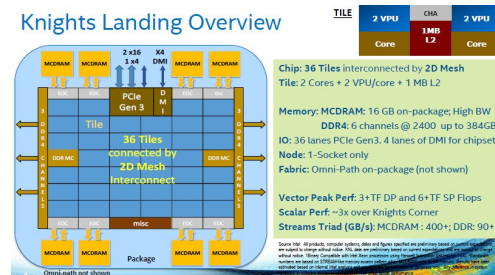
1. Based on the highest frequency. 2. Peak 100% for integrated fabric. Pricing shown is for only what is required. All additional I/O for integrated fabric version of these parts. Integrated fabric parts available in October.

Penn ESE532 Fall 2019 -- DeHon

7

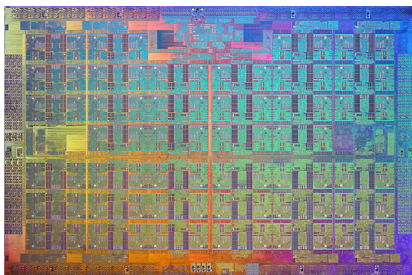
Intel Knights Landing

Knights Landing Overview



<https://www.nextplatform.com/2016/06/20/intel-knights-landing-yields-big-bang-buck-jump/>
 Penn ESE532 Fall 2019 -- DeHon [Intel, Micro 2016]

Knights Landing Xeon Phi



Penn ESE532 Fall 2019 -- DeHon

[Intel, Micro 2016]

9

What's happening?

- Fabricated chip has 76 cores
- Not expect all to work
- Selling based on functional cores
 - 72, 68, 64
- Charge premium for high core counts
 - Don't yield as often, people pay more
- **Do see design to accommodate defects**
<https://www.nextplatform.com/2016/08/22/intel-tweaking-xeon-phi-deep-learning/>

Penn ESE532 Fall 2019 -- DeHon

[Intel, Micro 2016]

10

Warmup Discussion

- Where else do we guard against defects today?
 - Where do we accept imperfection today?

Penn ESE532 Fall 2019 -- DeHon

11

Motivation: Probabilities

- Given:
 - N objects
 - P_g yield probability
- What's the probability for yield of composite system of N items? [Preclass 2]
 - Assume iid (independent, identically distributed) faults
 - $P(N \text{ items good}) = (P_g)^N$

Penn ESE532 Fall 2019 -- DeHon

12

Probabilities

- $P_{\text{all_good}}(N) = (P_g)^N$
- $P = 0.999999$

N	$P_{\text{all_good}}(N)$
10^4	
10^5	
10^6	
10^7	

Probabilities

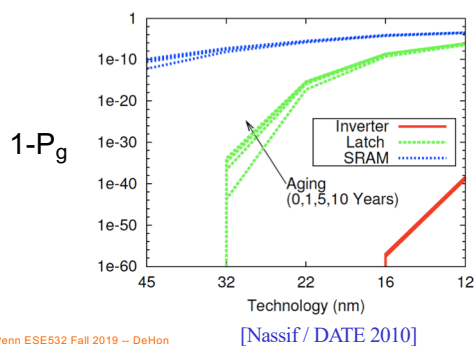
- $P_{\text{all_good}}(N) = (P_g)^N$
- $P = 0.999999$

N	$P_{\text{all_good}}(N)$
10^4	0.99
10^5	0.90
10^6	0.37
10^7	0.000045

Simple Implications

- As N gets large
 - must either increase reliability
 - ...or start tolerating failures
 - N
 - memory bits
 - disk sectors
 - wires
 - transmitted data bits
 - processors
 - transistors
 - molecules
- As devices get **smaller**, failure rates increase
chemists think $P=0.95$ is good
- As devices get **faster**, failure rate increases

Failure Rate Increases



Quality Required for Perfection?

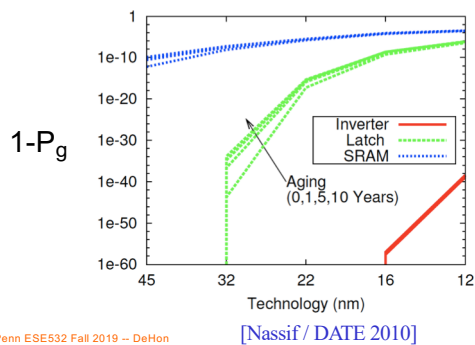
- How high must P_g be to achieve 90% yield on a collection of 10^{11} devices?

[preclass 4]

$$(P_g)^{10^{11}} > 0.9$$

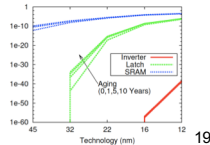
$$P_g > 1 - 10^{-12}$$

Failure Rate Increases



Challenge

- Feature size scales down (S)
- Capacity (area) \rightarrow increases ($1/S^2$)
 - N increase
- Reliability per device goes down
 - P_g decrease
- $P(N \text{ items good}) = (P_g)^N$



Defining Problems

Three Problems

- 1. Defects:** Manufacturing imperfection
 - Occur before operation; persistent
 - Shorts, breaks, bad contact
- 2. Transient Faults:**
 - Occur during operation; transient
 - node X value flips: crosstalk, ionizing particles, bad timing, tunneling, thermal noise
- 3. Lifetime “wear” defects**
 - Parts become bad during operational lifetime
 - Fatigue, electromigration, burnout....
 - ...slower
 - NBTI, Hot Carrier Injection

In a Nut-shell...

Shekhar Bokar
Intel Fellow
Micro37 (Dec.2004)



- 100 BT integration capacity
- 20 BT unusable (variations)
- 10 BT will fail over time
- Intermittent failures

Yet, deliver high performance in the power & cost envelope

First Step to Recover

Admit you have a problem
(observe that there is a failure)

Detection

- How do we determine if something wrong?
 - Some things easy
 - ...won't start
 - Others tricky
 - ...one **and** gate computes False & True \rightarrow True
- Observability
 - can see effect of problem
 - some way of telling if defect/fault present

Detection

- Coding
 - space of legal values \ll space of all values
 - should only see legal
 - e.g. parity, ECC (Error Correcting Codes)
- Explicit test (defects, recurring faults)
 - ATPG = Automatic Test Pattern Generation
 - Signature/BIST=Built-In Self-Test
 - POST = Power On Self-Test
- Direct/special access
 - test ports, scan paths

Penn ESE532 Fall 2019 -- DeHon

25

Coping with defects/faults?

- **Key idea:** **redundancy**
- Detection:
 - Use redundancy to detect error
- Mitigating: use redundant hardware
 - Use spare elements in place of faulty elements (defects)
 - Compute multiple times so can discard faulty result (faults)

Penn ESE532 Fall 2019 -- DeHon

26

Defect Tolerance

Penn ESE532 Fall 2019 -- DeHon

27

Three Problems

1. **Defects:** Manufacturing imperfection
 - Occur before operation; persistent
 - Shorts, breaks, bad contact
2. **Transient Faults:**
 - Occur during operation; transient
 - node X value flips: crosstalk, ionizing particles, bad timing, tunneling, thermal noise
3. **Lifetime “wear” defects**
 - Parts become bad during operational lifetime
 - Fatigue, electromigration, burnout...
 - ...slower
 - NBTI, Hot Carrier Injection

Penn ESE532 Fall 2019 -- DeHon

28

Two Models

- Disk Drives (defect map)
- Memory Chips (perfect chip)

Penn ESE532 Fall 2019 -- DeHon

29

Disk Drives

- Expose defects to software
 - software model expects defects
 - Create table of good (bad) sectors
 - manages by masking out in software
 - (at the OS level)
 - Never allocate a bad sector to a task or file
 - yielded capacity varies

Penn ESE532 Fall 2019 -- DeHon

30

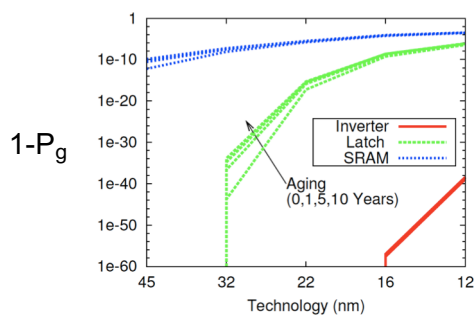
Memory Chips

- Provide model in **hardware** of perfect chip
- Model of perfect memory at capacity X
- Use redundancy in hardware to provide perfect model
- Yielded capacity fixed
 - discard part if not achieve

Example: Memory

- Correct memory:
 - N slots
 - each slot reliably stores last value written
- Millions, billions, etc. of bits...
 - have to get them all right?

Failure Rate Increases

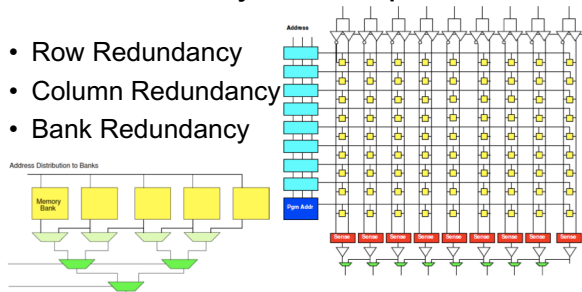


Memory Defect Tolerance

- Idea:
 - few bits may fail
 - provide more raw bits
 - configure so yield what looks like a perfect memory of specified size

Memory Techniques

- Row Redundancy
- Column Redundancy
- Bank Redundancy



Yield M of N

- Preclass 5: Probability of yielding 3 of 5 things?
 - Symbolic?
 - Numerical for $P_g=0.9$?

Yield M of N

- $P(M \text{ of } N) = P(\text{yield } N)$
 - + $(N \text{ choose } N-1) P(\text{exactly } N-1)$
 - + $(N \text{ choose } N-2) P(\text{exactly } N-2) \dots$
 - + $(N \text{ choose } N-M) P(\text{exactly } N-M) \dots$
- [think binomial coefficients]

M of 5 example

- $1 \cdot P^5 + 5 \cdot P^4(1-P)^1 + 10P^3(1-P)^2 + 10P^2(1-P)^3 + 5P^1(1-P)^4 + 1 \cdot (1-P)^5$
- Consider $P=0.9$
 - $1 \cdot P^5$ 0.59 M=5 $P(\text{sys})=0.59$
 - $5 \cdot P^4(1-P)^1$ 0.33 M=4 $P(\text{sys})=0.92$
 - $10P^3(1-P)^2$ 0.07 M=3 $P(\text{sys})=0.99$
 - $10P^2(1-P)^3$ 0.008 Can achieve higher system yield than individual components!
 - $5P^1(1-P)^4$ 0.00045
 - $1 \cdot (1-P)^5$ 0.00001

Possible Yield of 76 cores@ P=0.9

Processors Yield	Prob Exact	Prob at least
76	0.001	0.001
75	0.004	0.005
74	0.016	0.020
73	0.041	0.061
72	0.079	0.140
71	0.119	0.259
70	0.148	0.407
69	0.156	0.562
68	0.141	0.704
67	0.112	0.816
66	0.079	0.895
65	0.050	0.945
64	0.028	0.973

Possible Yield of 76 cores@ P=0.9

Processors Yield	Prob at least
76	0.001
75	0.005
74	0.020
73	0.061
72	0.140
71	0.259
70	0.407
69	0.562
68	0.704
67	0.816
66	0.895
65	0.945
64	0.973

Out of 100 chips, how many?

Sell with 72:
Sell with 68:
Sell with 64:
Discard:

Intel Xeon Phi Pricing

	CORES	GHZ	MEMORY	FABRIC	DDR4	POWER ²	RECOMMENDED CUSTOMER PRICING ¹
7290¹ Best Performance/Node	72	1.5	16GB 7.2 GT/s	Yes	384GB 2400 MHz	245W	\$6254
7250 Best Performance/Watt	68	1.4	16GB 7.2 GT/s	Yes	384GB 2400 MHz	215W	\$4876
7230 Best Memory Bandwidth/Core	64	1.3	16GB 7.2 GT/s	Yes	384GB 2400 MHz	215W	\$3710
7210 Best Value	64	1.3	16GB 6.4 GT/s	Yes	384GB 2133 MHz	215W	\$2438

Repairable Area

- Not all area in a RAM is repairable
 - memory bits spare-able
 - io, power, ground, control not redundant

Repairable Area

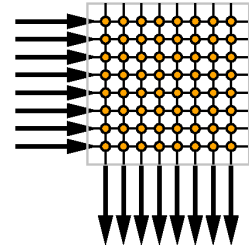
- $P(\text{yield}) = P(\text{non-repair}) * P(\text{repair})$
- $P(\text{non-repair}) = P^{N_{nr}}$
 - $N_{nr} \ll N_{\text{total}}$
 - $P > P_{\text{repair}}$
 - e.g. use coarser feature size
 - Differential reliability
- $P(\text{repair}) \sim P(\text{yield } M \text{ of } N)$

Penn ESE532 Fall 2019 -- DeHon

43

Consider a Crossbar

- Allows us to connect any of N things to each other
- E.g.
 - N processors
 - N memories
 - N/2 processors
 - + N/2 memories

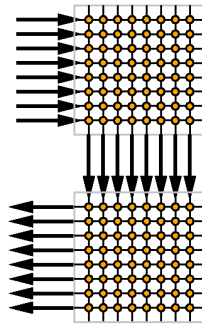


Penn ESE532 Fall 2019 -- DeHon

44

Crossbar Buses and Defects

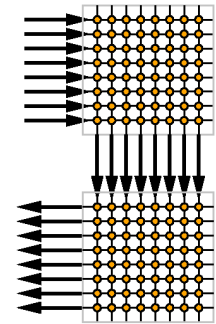
- Two crossbar multibus
- Wires may fail
- Switches may fail
- How tolerate
 - Wire failures between crossbars?
 - Switch failures?



Penn ESE532 Fall 2019 -- DeHon

Crossbar Buses and Defects

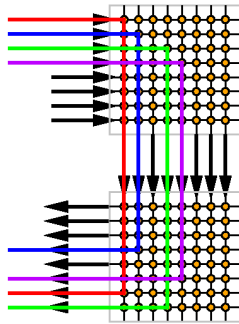
- Two crossbars
- Wires may fail
- Switches may fail
- Provide more wires
 - Any wire fault avoidable
 - M choose N



Penn ESE532 Fall 2019 -- DeHon

Crossbar Buses and Defects

- Two crossbars
- Wires may fail
- Switches may fail
- Provide more wires
 - Any wire fault avoidable
 - M choose N

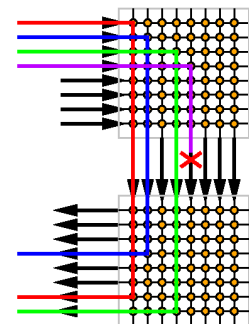


Penn ESE532 Fall 2019 -- DeHon

47

Crossbar Buses and Faults

- Two crossbars
- Wires may fail
- Switches may fail
- Provide more wires
 - Any wire fault avoidable
 - M choose N

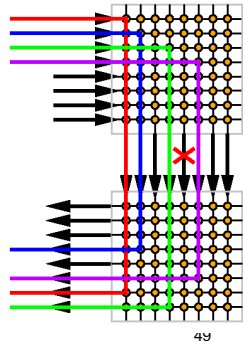


Penn ESE532 Fall 2019 -- DeHon

48

Crossbar Buses and Faults

- Two crossbars
- Wires may fail
- Switches may fail
- Provide more wires
 - Any wire fault avoidable
 - M choose N
 - Same idea

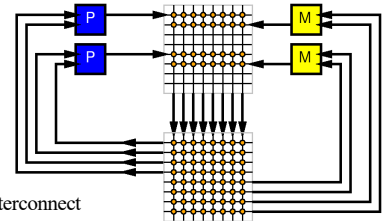


Penn ESE532 Fall 2019 -- DeHon

49

Simple System

- P Processors
- M Memories
- Wires

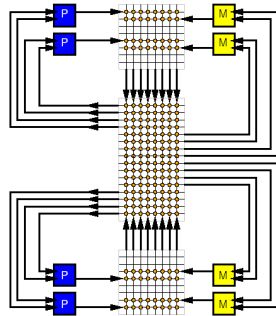


Memory, Compute, Interconnect

Penn ESE532 Fall 2019 -- DeHon

Simple System w/ Spares

- P Processors
- M Memories
- Wires
- Provide spare
 - Processors
 - Memories
 - Wires

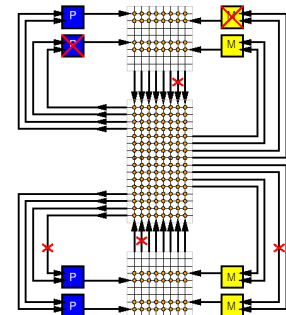


Penn ESE532 Fall 2019 -- DeHon

51

Simple System w/ Defects

- P Processors
- M Memories
- Wires
- Provide spare
 - Processors
 - Memories
 - Wires
- ...and defects

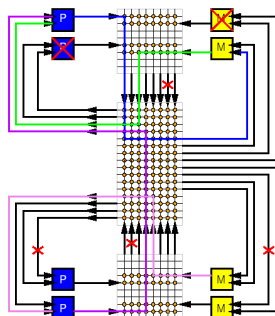


Penn ESE532 Fall 2019 -- DeHon

52

Simple System Repaired

- P Processors
- M Memories
- Wires
- Provide spare
 - Processors
 - Memories
 - Wires
- Use crossbar to switch together good processor and memories

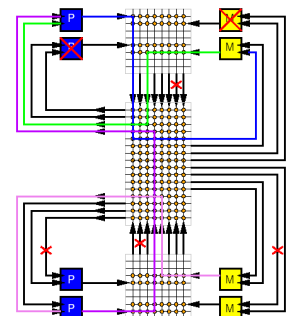


Penn ESE532 Fall 2019 -- DeHon

53

Simple System Repaired

- What are the costs?
 - Area
 - Energy
 - Delay



Penn ESE532 Fall 2019 -- DeHon

54

In Practice

- Crossbars are inefficient
- Use switching networks with
 - Locality
 - Segmentation
- ...but basic idea for sparing is the same



55

Penn ESE532 Fall 2019 -- DeHon

FPGAs

56

Penn ESE532 Fall 2019 -- DeHon

Modern FPGA

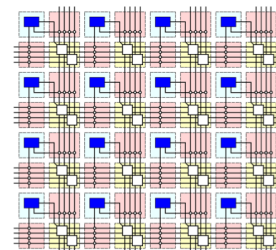
- Has 10,000 to millions of LUTs
- Hundreds to thousands of
 - Memory banks
 - Multipliers
- Reconfigurable interconnect

57

Penn ESE532 Fall 2019 -- DeHon

ZU3EG (Ultra96)

- 6-LUTs: 70,560
- DSP Blocks: 360
 - 18x27 multiply
 - 48b accumulate
- BRAMs: 216
 - 36Kb
 - Dual port
 - Up to 72b wide (512x72)



58

Penn ESE532 Fall 2019 -- DeHon

Modern FPGA

- Has 10,000 to millions of gates
- Hundreds to thousands of
 - Memory banks
 - Multipliers
- Reconfigurable interconnect
- If a few resources don't work
 - avoid them

59

Penn ESE532 Fall 2019 -- DeHon

Granularity

- Consider two cases
 - Knight's Bridge with 76 64b processors
 - FPGA with 1 Million 6-LUTs
- 10 defects (bad transistors)
- How much capacity lost?
 - [worst-case, as percentage]
 - Knight's Bridge?
 - 1M 6-LUT FPGA?

60

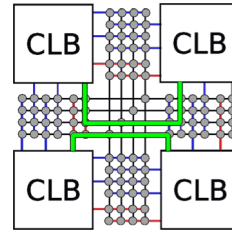
Penn ESE532 Fall 2019 -- DeHon

Observe

- Finer granularity sparing
 - Lose fewer resources per defect
 - Losing processor vs. losing 6-LUT
 - Losing memory row vs. losing entire memory bank
 - Pay more for reconfiguration
- Finer grained designs
 - Tolerate higher defect rates

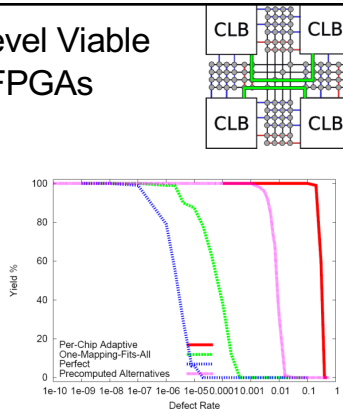
Interconnect Defects

- Route around interconnect defects



Defect-Level Viable with FPGAs

- Fine-grained repair
- Avoiding routing defects
 - Tolerates >20% switch defects

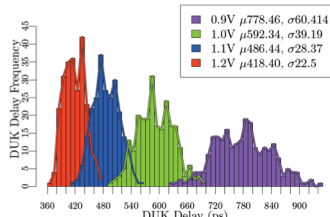


FPGAs Variation and Energy

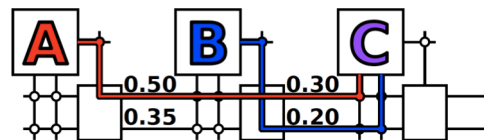
(if time permits)

Variation

- Modern ICs are like Snowflakes
 - Each one is different

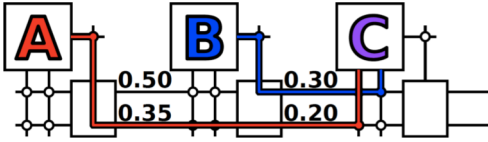


Variation Challenge



- Use of high V_{th} resource forces high supply voltage (V_{dd}) to meet timing requirement
- Delay: CV/I and I goes as $(V_{dd}-V_{th})^2$

Component-Specific



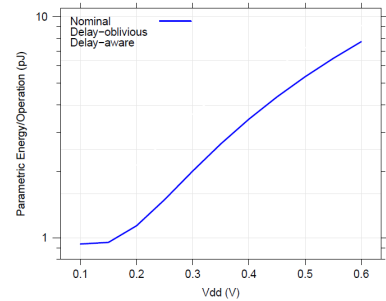
- Use defect idea to avoid high V_{th} resource
- Allow lower supply voltage (V_{dd}) to meet timing requirement
- Delay: CV/I and I goes as $(V_{dd}-V_{th})^2$

Penn ESE532 Fall 2019 -- DeHon

67

Energy vs V_{dd}

- Nominal uses minimum size



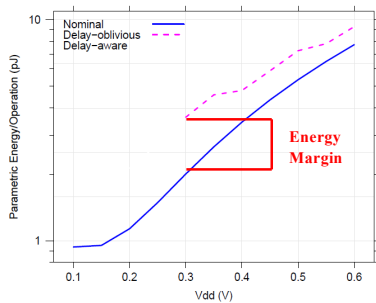
Penn ESE532 Fall 2019 -- DeHon

[Mehta, FPGA 2012]

68

Energy vs V_{dd} (des)

- Nominal uses minimum size
- Oblivious must use larger
 - Higher C, Area



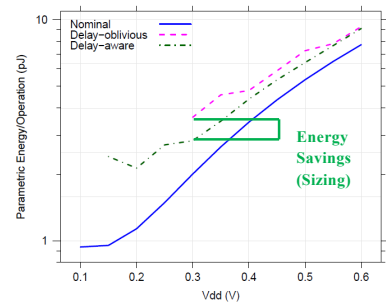
Penn ESE532 Fall 2019 -- DeHon

[Mehta, FPGA 2012]

69

Energy vs V_{dd} (des)

- Nominal uses minimum size
- Oblivious must use larger
 - Higher C, Area
- Delay-aware routing reduces energy margins
 1. Smaller sizes



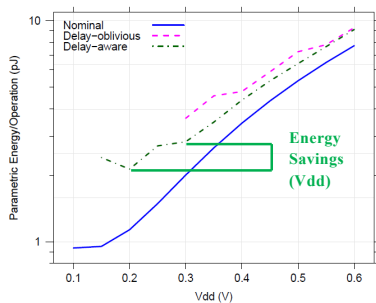
Penn ESE532 Fall 2019 -- DeHon

[Mehta, FPGA 2012]

70

Energy vs V_{dd} (des)

- Nominal uses minimum size
- Delay-aware routing reduces energy margins
 1. Smaller sizes
 2. Lower voltages



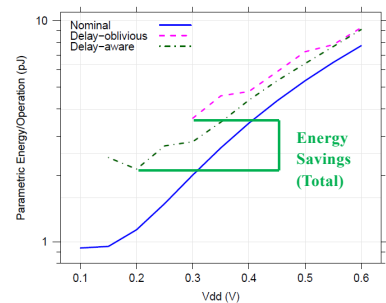
Penn ESE532 Fall 2019 -- DeHon

[Mehta, FPGA 2012]

71

Energy vs V_{dd} (des)

- Nominal uses minimum size
- Delay-aware routing reduces energy margins
 1. Smaller sizes
 2. Lower voltages
 3. Less Leakage



Penn ESE532 Fall 2019 -- DeHon

[Mehta, FPGA 2012]

72

Big Ideas

- At small feature sizes, not viable to demand perfect fabrication of billions of transistors on a chip
- Modern ICs are like snowflakes
 - Everyone is different, changes over time
- Reconfiguration allows repair
 - Finer grain → higher defect rates
 - Tolerate variation → lower energy

Admin

- Project due Friday
 - Reminder turnin
 - Report
 - Code
 - Binaries: Bitstream, elf, decoder
- Wrapup lecture on Monday by Eric
 - This was last lecture from André
 - Plan to return boards, cables, cards
 - Watch piazza for more details