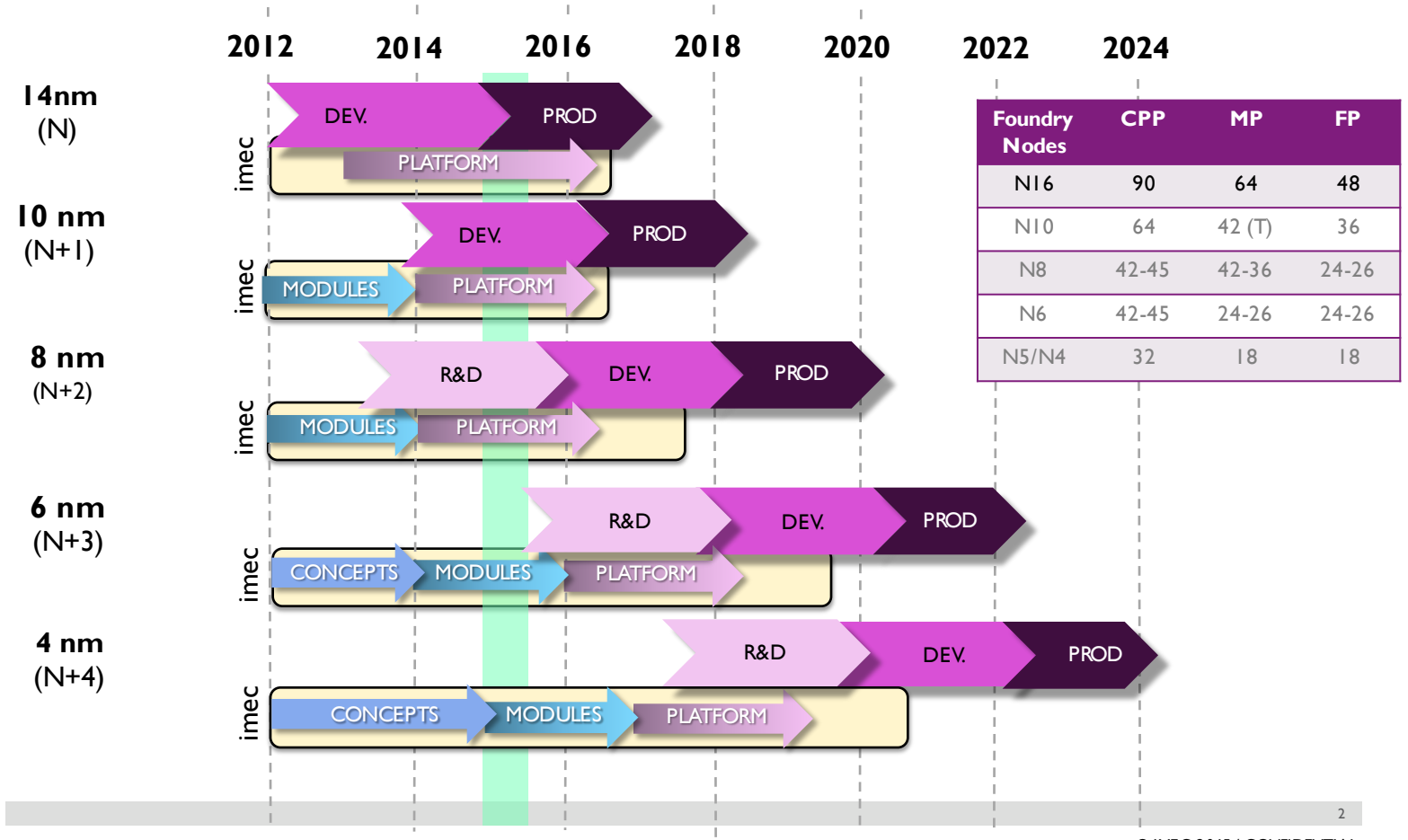
??nm node

Holistic device exploration for 7nm node

PRAVEEN RAGHAVAN

imec

LIKELY FOUNDRY LOGIC TECHNOLOGY TIMELINE



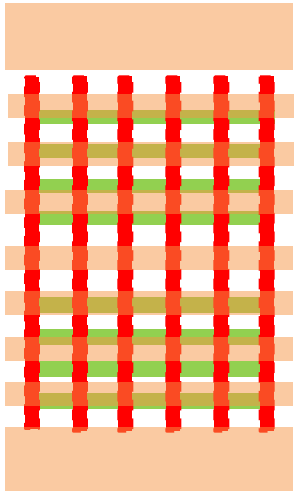
Foundry Nodes	CPP	MP	FP
N16	90	64	48
N10	64	42 (T)	36
N8	42-45	42-36	24-26
N6	42-45	24-26	24-26
N5/N4	32	18	18

Call me stupid:
What defines a node?

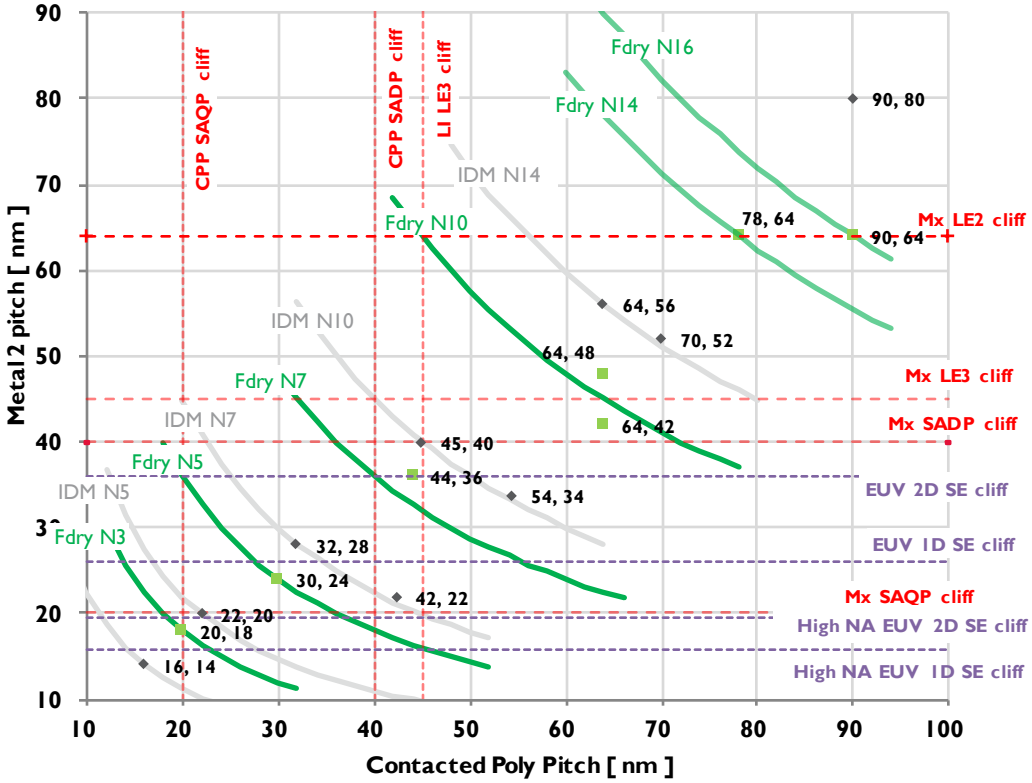
In the 14nm node, there is nothing that is 14nm (perhaps the gate spacer?)

So what is a node name?

- A combination of Metal Pitch and Poly Pitch
- A combination of secondary rules
  - Fin termination, metal tip to tip, special constructs
- Technology readiness (when can be yield)
- A ton of marketing to capture the market! Yes!

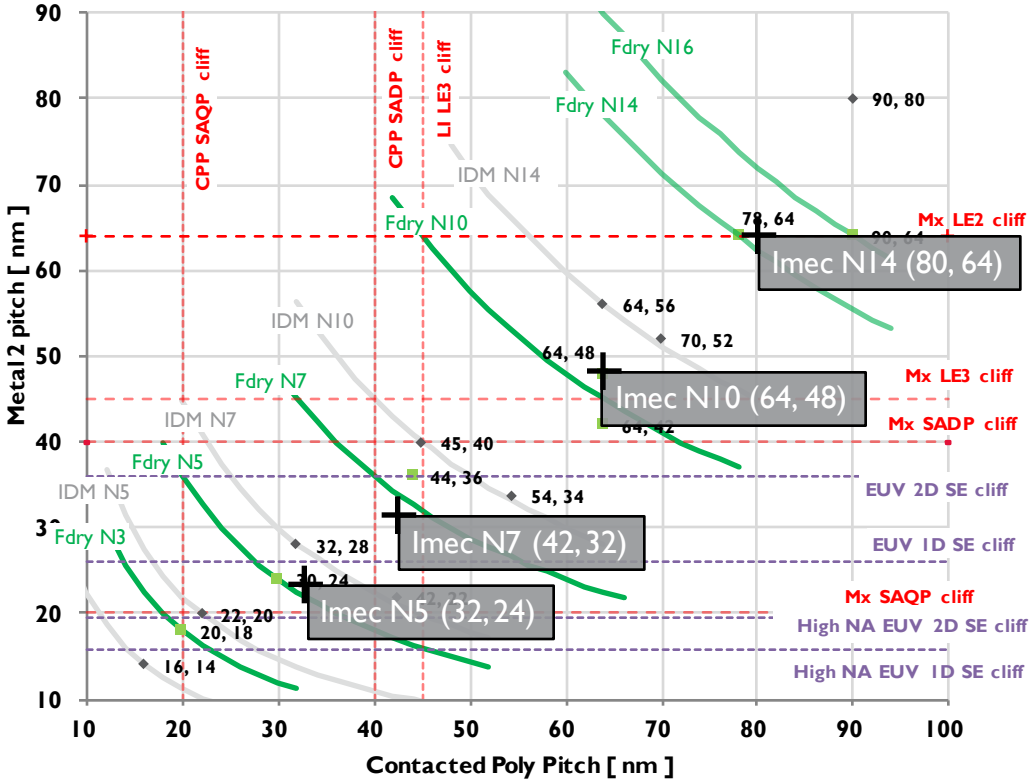


LOGIC SCALING LANDSCAPE



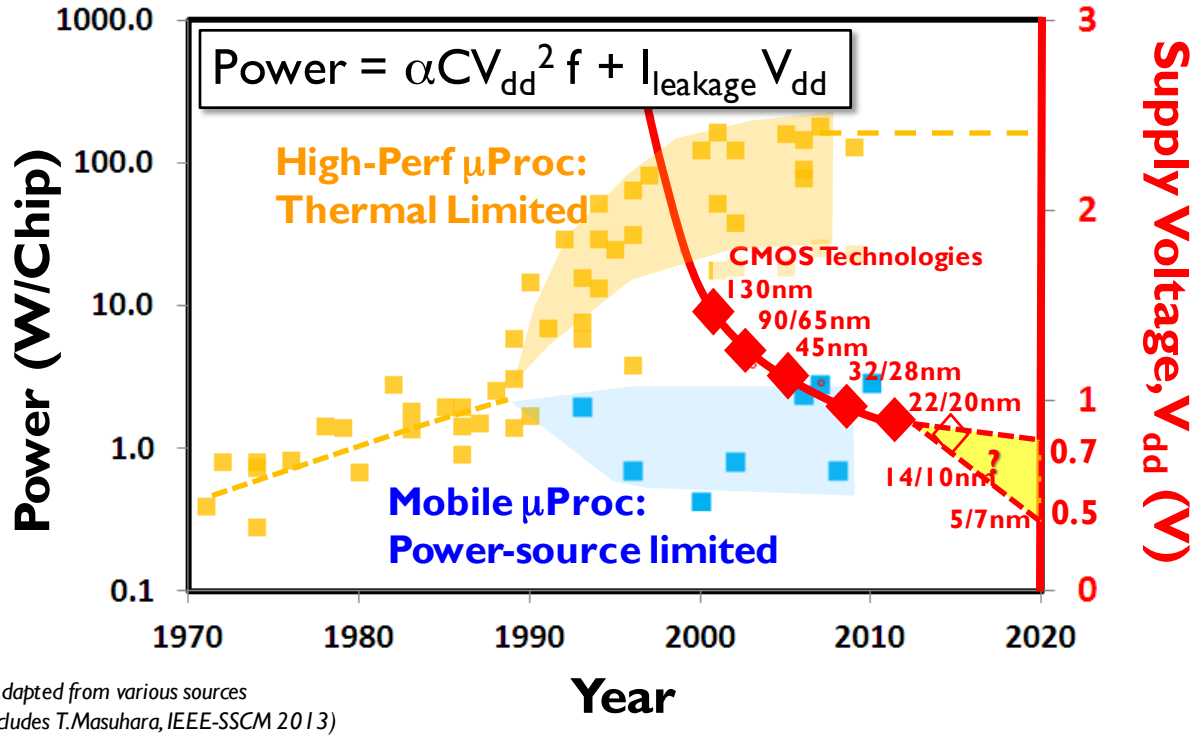
- MP x CPP determines logic cell area
- Choices of MP, CPP
  - Patterning cliffs, cost
  - Layout style (1D, 2D)
  - Electrical performance
    - Reduced CPP forces trade-off between device electrostatics (gate-length) and parasitics (spacer, contact sizes)
    - Reduced MP leads to steep increase in BEOL resistance

LOGIC SCALING LANDSCAPE



- MP x CPP determines logic cell area
- Choices of MP, CPP
  - Patterning cliffs, cost
  - Layout style (1D, 2D)
  - Electrical performance
    - Reduced CPP forces trade-off between device electrostatics (gate-length) and parasitics (spacer, contact sizes)
    - Reduced MP leads to steep increase in BEOL resistance

POWER CRISIS DUE TO VDD SCALABILITY



(Adapted from various sources
Includes T.Masuhara, IEEE-SSCM 2013)

- Higher-Perf & Mobile Processors are power limited
- V_{dd} scaling slowing
- **Device utilization diminishes**
  → "Dark Silicon" Syndrome?

SOC SCALING CHALLENGES

Need to reduce supply voltage to be able to effectively use the growing number of devices on a die and beat dark silicon syndrome.

- Requires a device with good electro-static control

	Happy scaling	Dark silicon
Area	x0.5	x0.5
Voltage	x0.7	x1
Capacitance	x0.7	x0.7
Frequency	x1.4	x1
Power CV^2f	x0.5	x0.7
W/mm2	x1	x1.4

When power exceeds thermal dissipation limit, cannot power up all transistors on the die: dark silicon.

SOC SCALING CHALLENGES

	Happy scaling	Dark silicon
Area	x0.5	x0.5
Voltage	x0.7	x1
Capacitance	x0.7	x0.7
Frequency	x1.4	x1
Power CV^2f	x0.5	x0.7
W/mm2	x1	x1.4

When power exceeds thermal dissipation limit, cannot power up all transistors on the die: dark silicon.

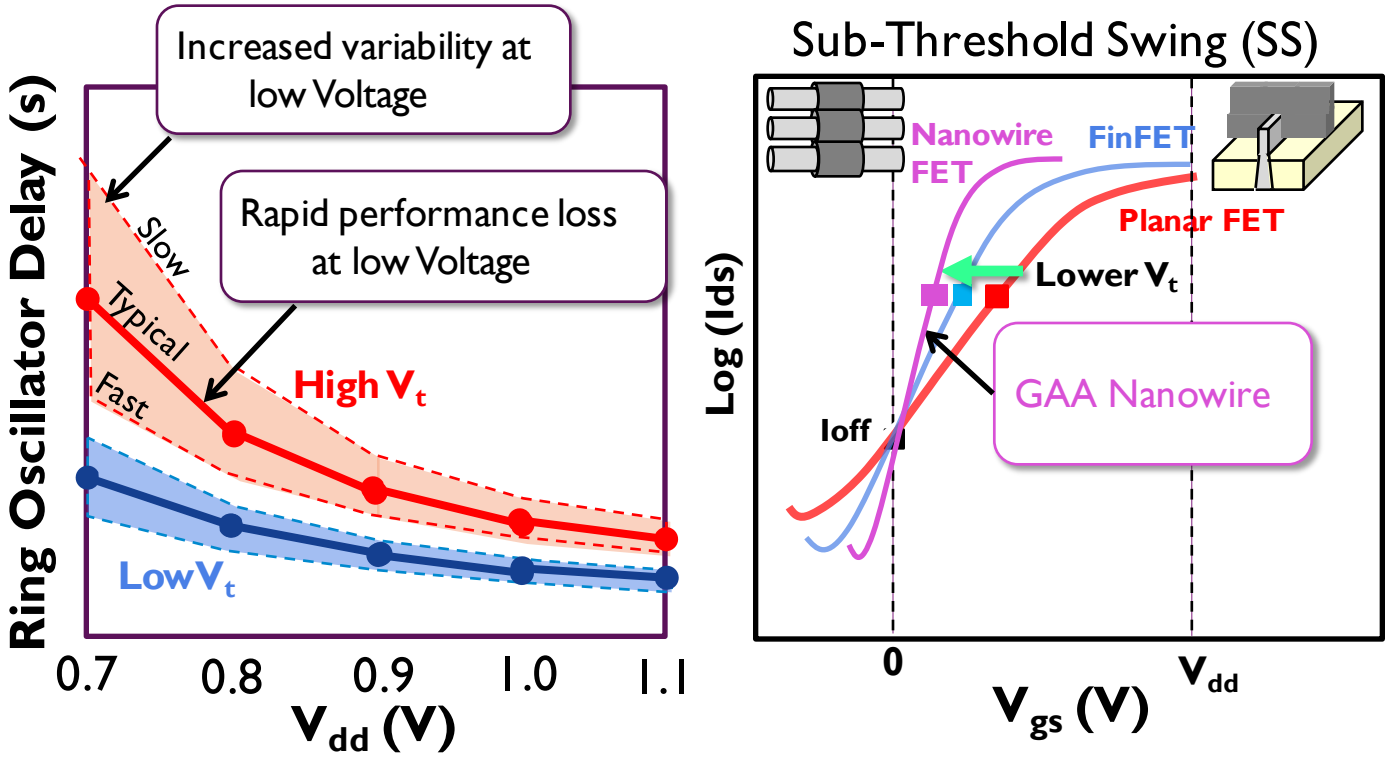
Need to reduce supply voltage to be able to effectively use the growing number of devices on a die and beat dark silicon syndrome.

- Requires a device with good electro-static control

Patterning complexity drives up node to node wafer cost increase beyond traditional 15%

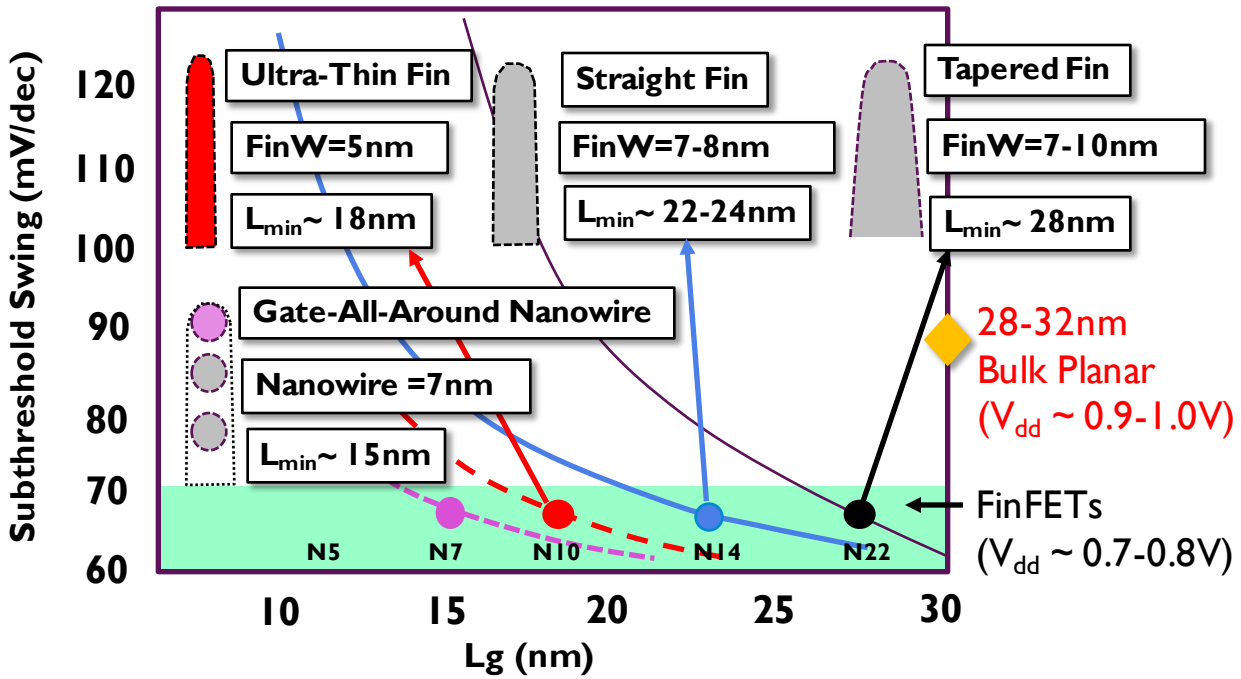
- Need to scale transistor density by more than 50% to maintain node-to-node transistor cost reduction of x0.6.

WHY ELECTRO-STATICS ARE KEY



- Circuit performance loss & variability limits V_{dd} scaling
- Transistors need to improve electrostatics, drive, & mitigate variability

DEVICE ARCHITECTURE IMPACTS ELECTROSTATICS



- FinFETs offered a Low-Voltage transistor option wrt bulk planar.
- To maintain electrostatics, simple FinFETs will hit limits

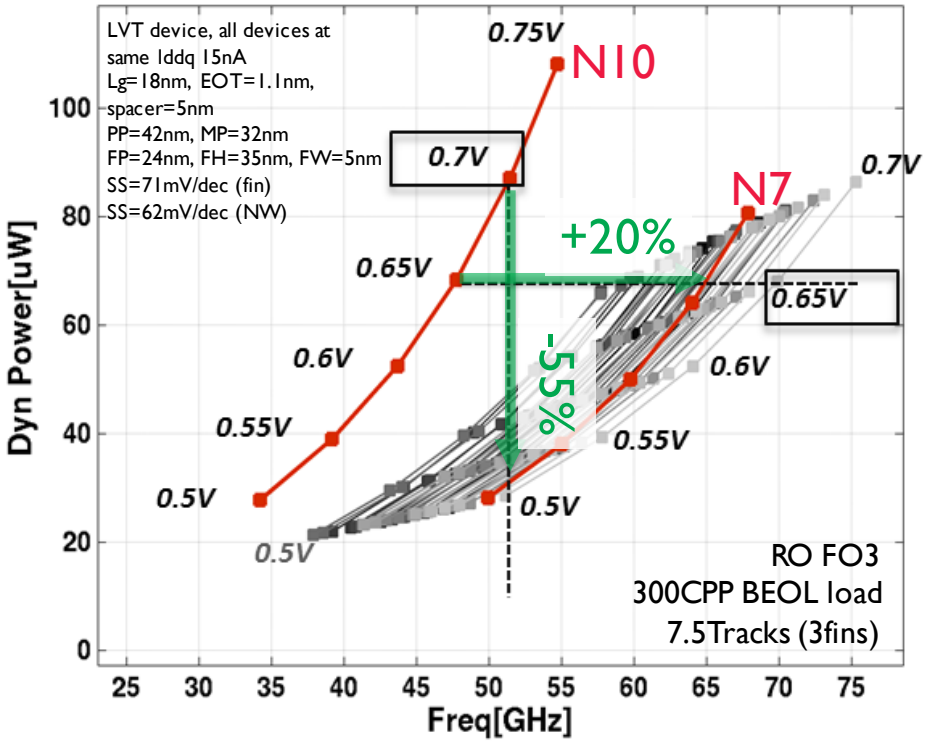
SOC SCALING REQUIREMENTS

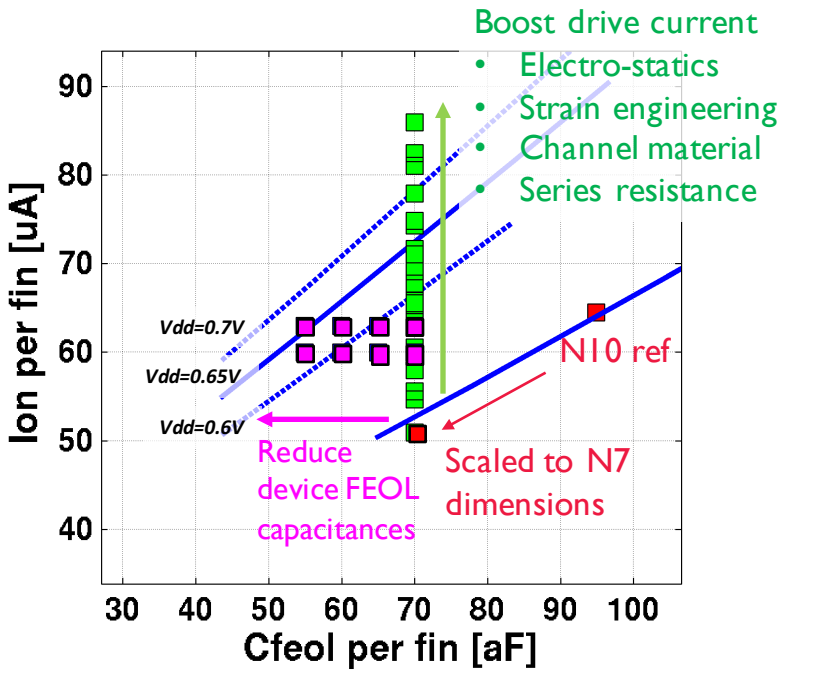
Need to reduce supply voltage to be able to effectively use the growing number of devices on a die (beat dark silicon syndrome).

Requires a device with good electro-static control (steep sub-threshold slope): FinFET outperform planar devices.

When scaling gate length to improve performance also need to scale fin-width to keep good electro-static behavior or introduce gate-all-around devices

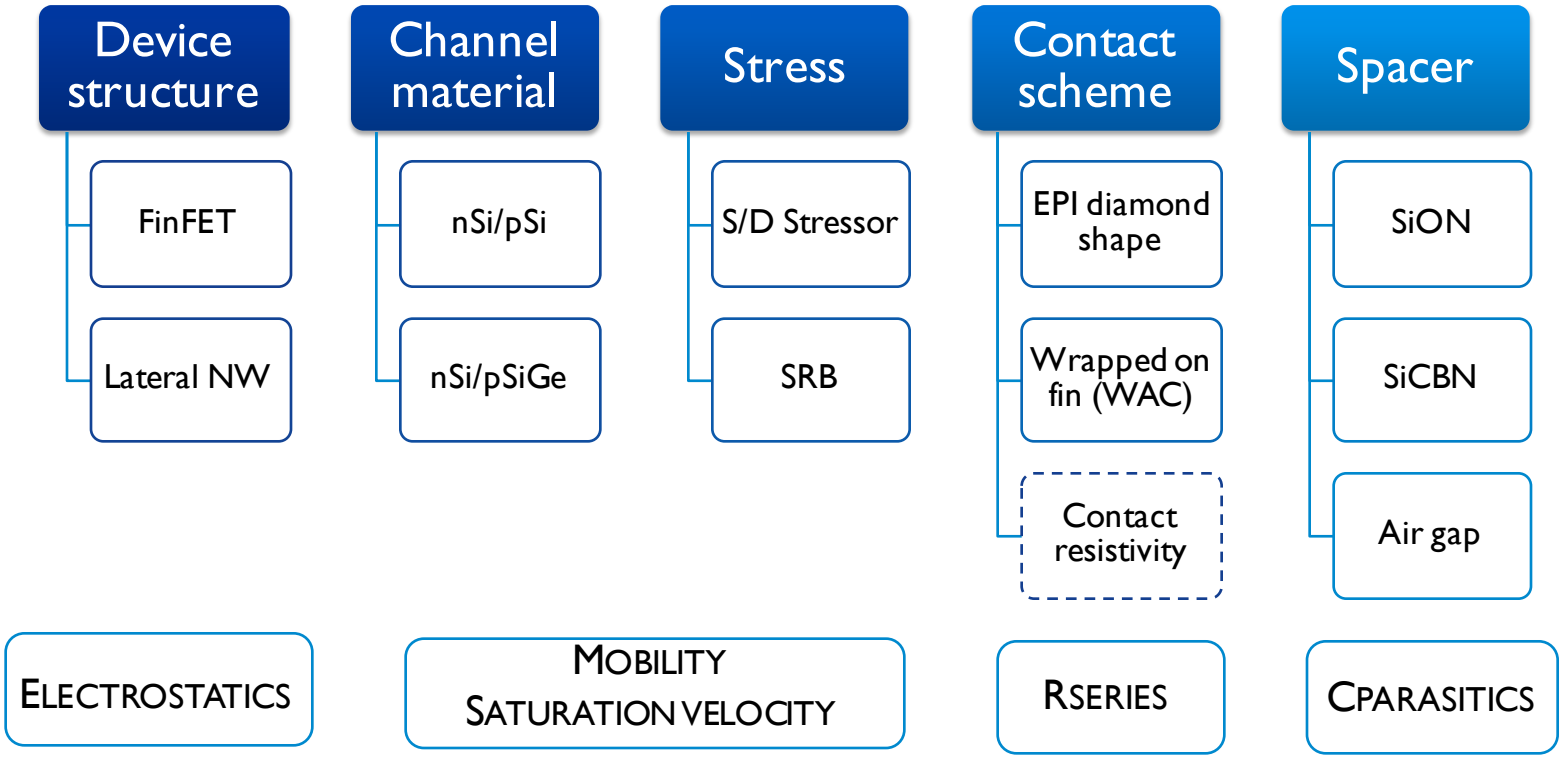
Device options for N7 target



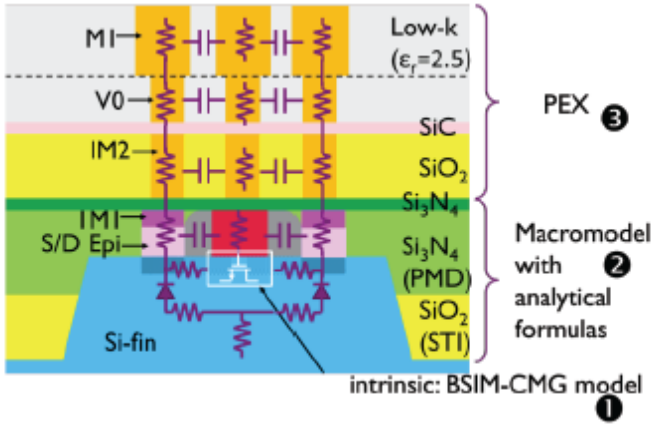


REMARK: HFIN, NFIN, increase both Cfeol and ION

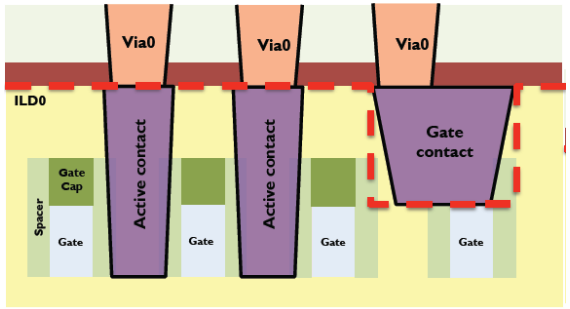
Performance elements for FEOL

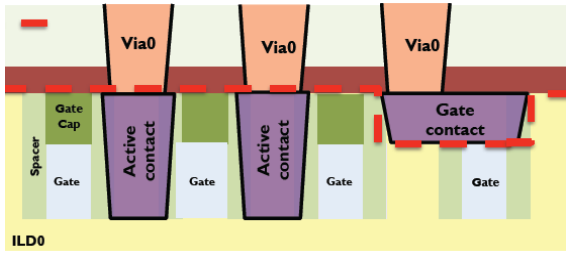


CMG – PARASITIC MACRO MODEL – PEX



- Everything above red dashed line is handled by PEX deck
- Everything below the red-line is taken care off by parasitic macro model.
- Parasitic capacitances crossing the red-line boundary are taken care of by PEX deck.





OUR BENCHMARK FOR TARGETING

Ring oscillator with INV in 9T or 7.5T version

Fanout 3

BEOL load of 300 CGP equivalent length of Mx

Performance elements for FEOL

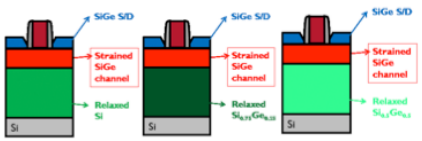
stressor	stress in channel
S/D EPI SiGe 50%	0.5 GPa
S/D EPI SiGe 80%	0.75 GPa
SRB SiGe 15%	1 GPa
SRB SiGe 21%	1.5 GPa

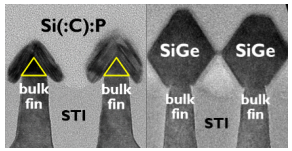
Channel material

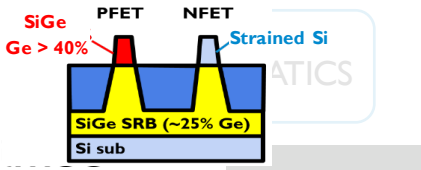
- nSi/pSi
- nSi/pSiGe

Stress

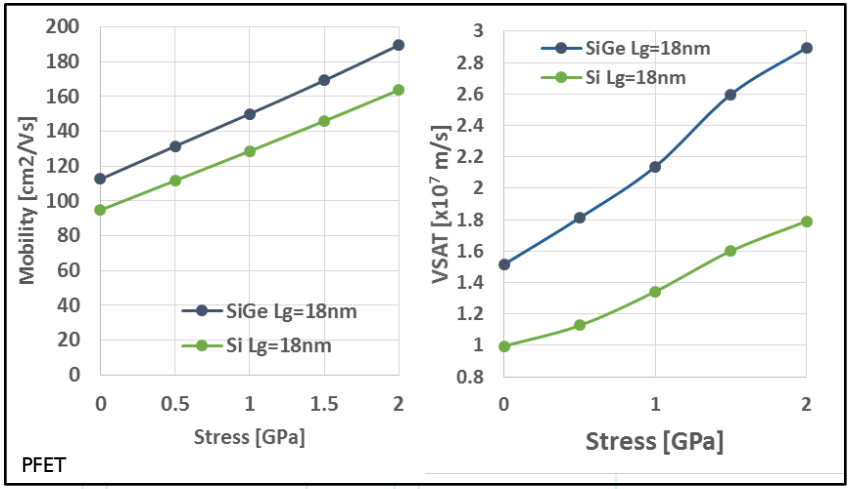
- S/D Stressor
- SRB

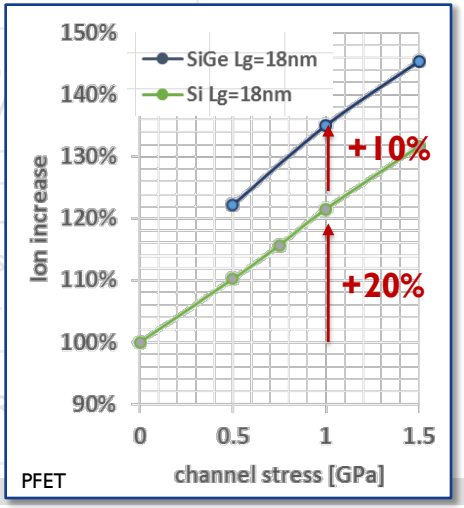




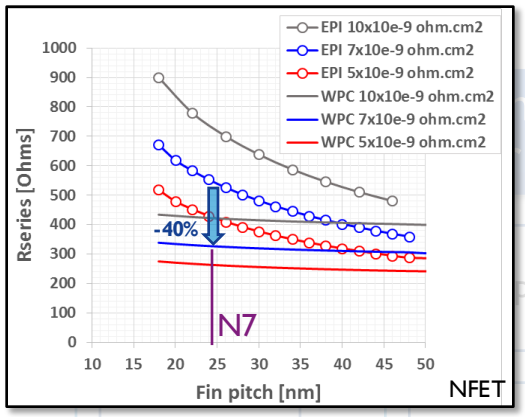


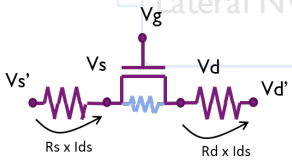
MOBILITY
SATURATION VELOCITY

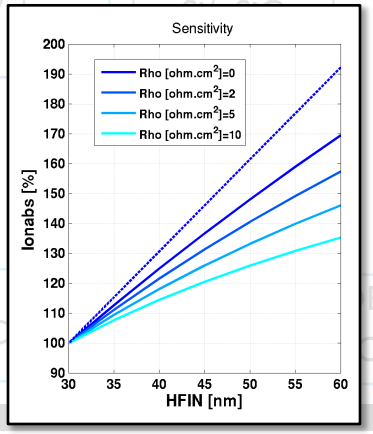


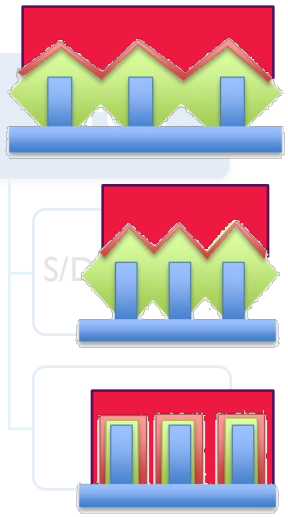


Performance elements for FEOL







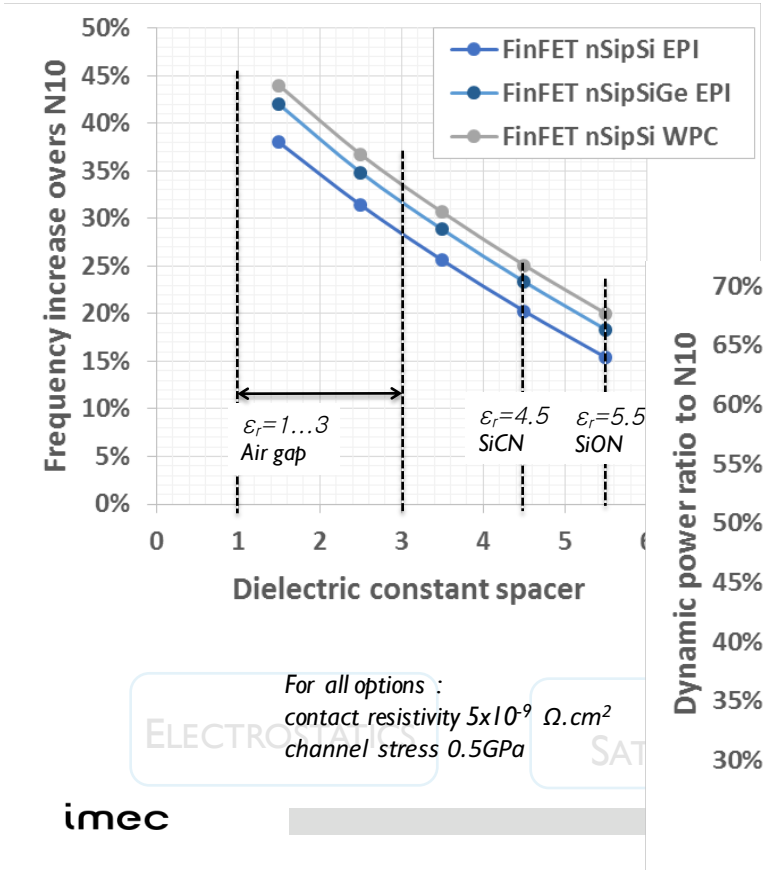


Contact scheme

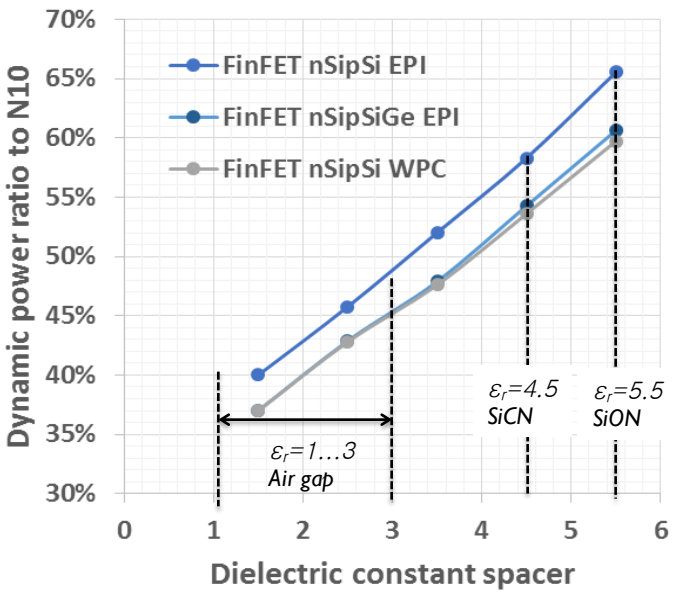
- EPI diamond shape — Scaling fin pitch quickly reduces contact area
- Wrapped on fin (WAC) — Large contact area and growing with fin height.
- Contact resistivity

RSERIES

Performance elements for FEOL



For all options :
contact resistivity 5x10^{-9} Ω.cm^2
channel stress 0.5GPa



Stress

Contact scheme

EPI diamond

Spacer

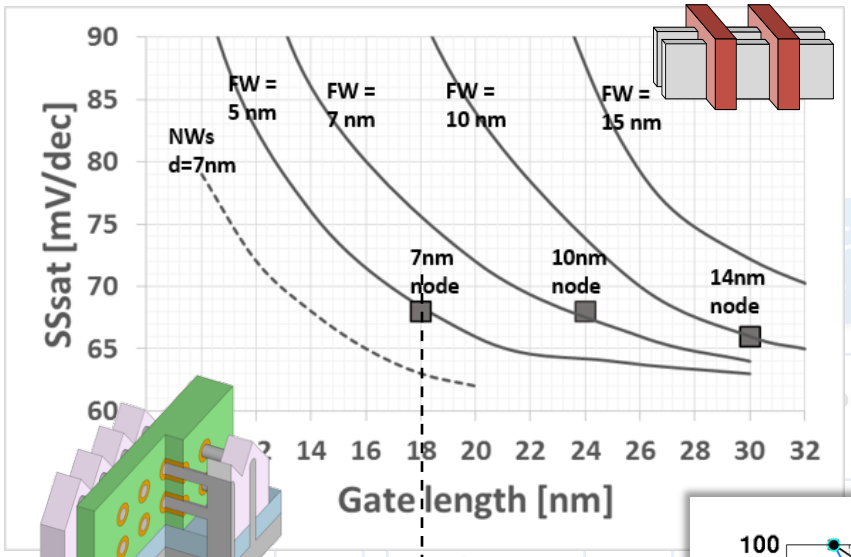
- SiON
- SiCBN
- Air gap

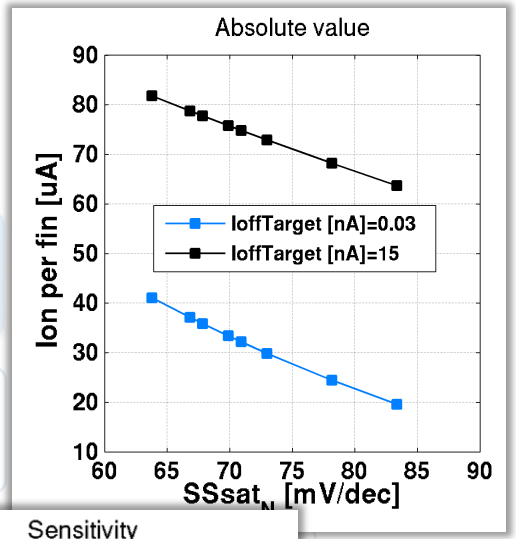
CPARASITICS

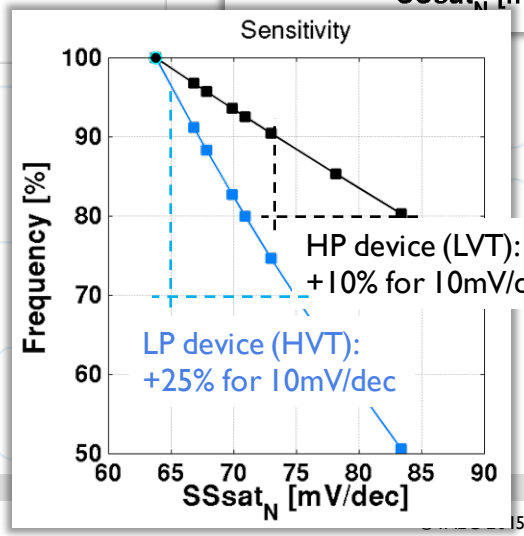
Performance

Device structure

- FinFET
- Lateral NW







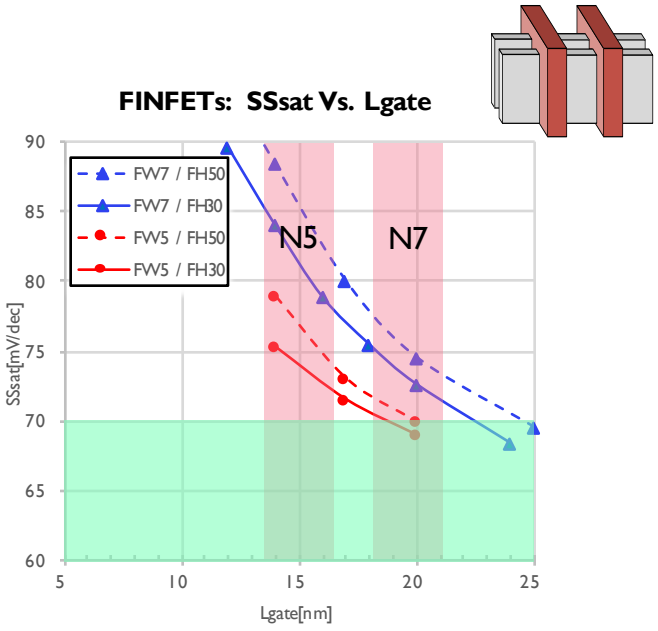
At Lg=18nm
Nanowires SS: 61 to 65 mV/dec
FinFETs SS: 68 to 72 mV/dec

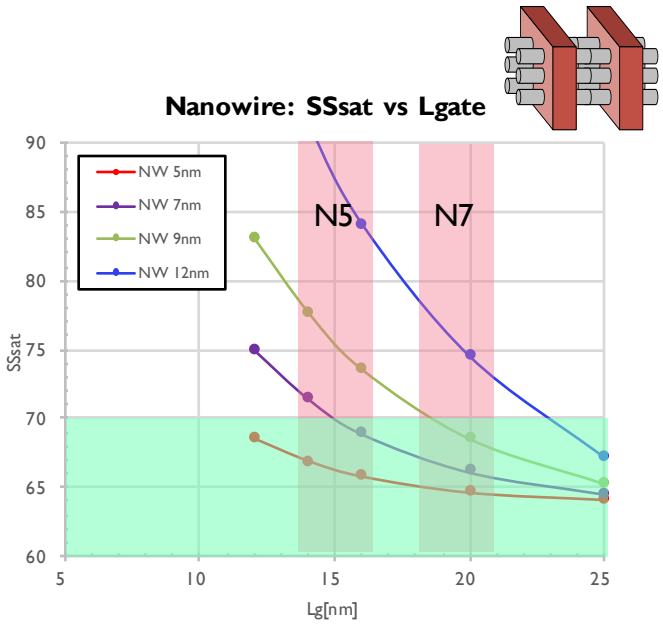
ELECTROSTATICS

MOBILITY

SATURATION VELOCITY

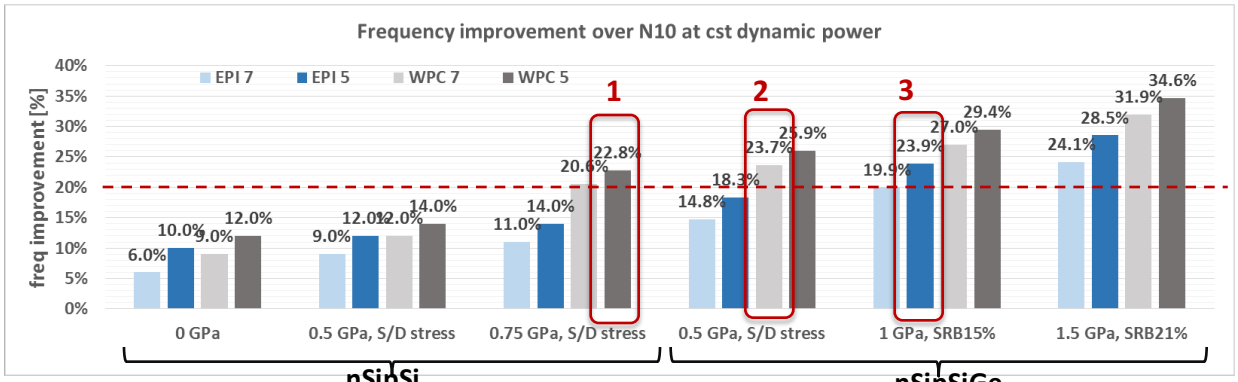
ELECTROSTATIC CLIFFS FOR DEVICE ARCHITECTURE

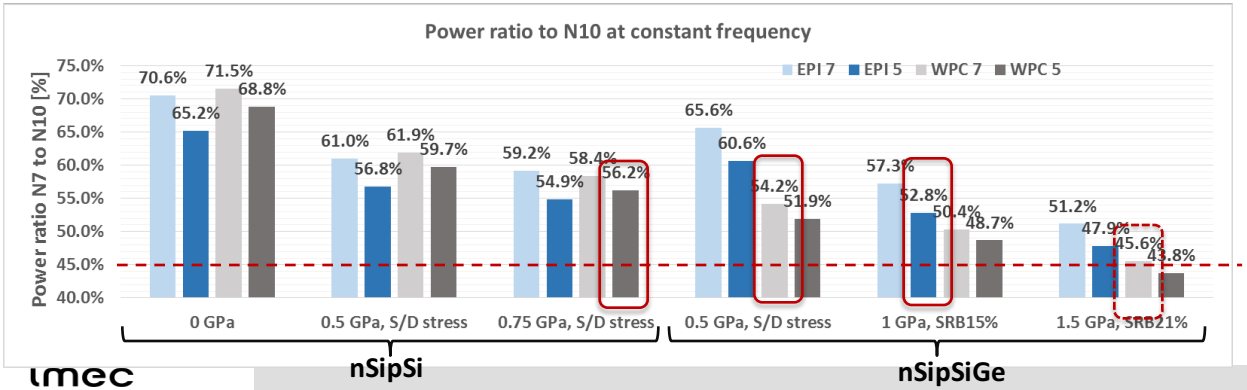




- Fin of width 5nm or nanowire of width 7nm option for N7/CPP-42nm
- Nanowire of width 5nm (or finwidth of 3nm) option for N5/CPP-32nm

N7 FINFET OPTIONS





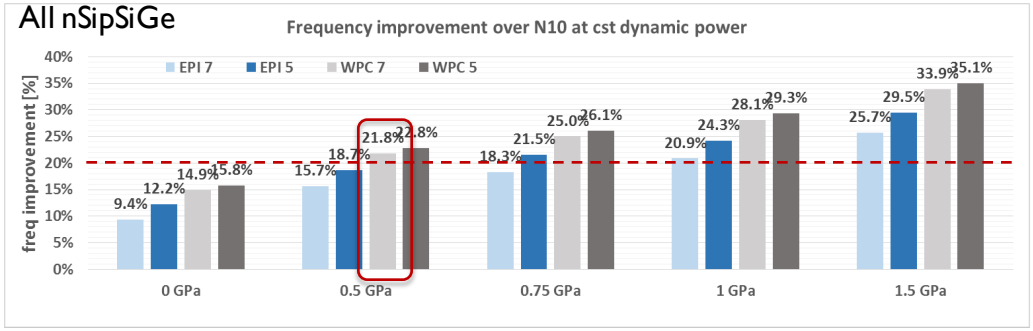
Minimum risk/ complexity options for this target speed

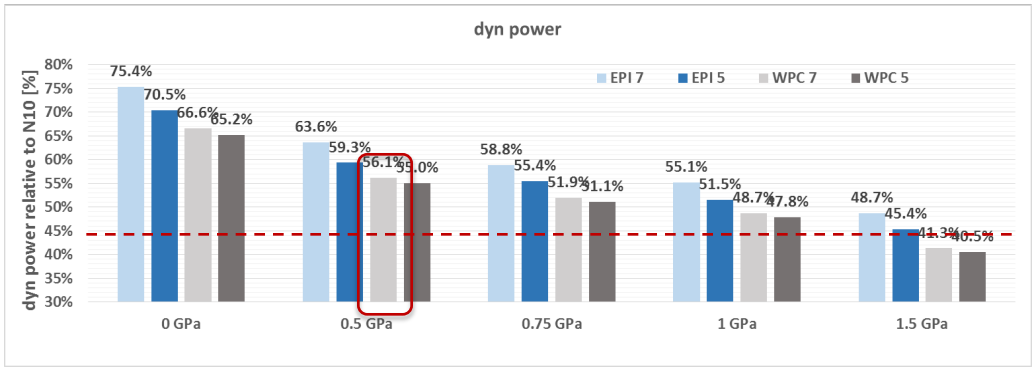
Corresponding dynamic power higher than target 45% from N10

RO FO3 INV
Wire length 300CGP
NFIN= 3
Vdd= 0.65V
Rbeol = 135 ohm/um
Cbeol = 0.16 fF/um

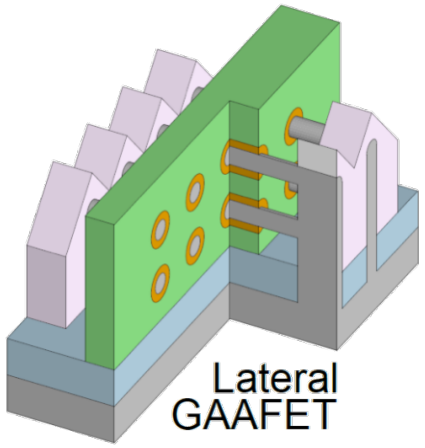
imec

N7 NANOWIRE OPTIONS





LNW options for N7 in similar performance range as FinFET



LOWER RISK OPTIONSTO MEET N7 TARGET 20% SPEED IMPROVEMENT

Device	Channel	Stress	Contact scheme	Contact resistivity [x10^{-9}Ω.cm^2]	Spacer dielectric constant	Sub slope	Ion per fin NFET	Ion per fin PFET	Freq. increase from N10 @cst power	Power ratio to N10 @cst speed
FinFET	nSi pSi	0.75GPa S/D stress	WPC	5	5.5	71 mV/dec	72.1 uA	69.8 uA	22.8%	58.2%
FinFET	nSi pSiGe	0.5GPa S/D stress	WPC	7	5.5	71 mV/dec	65.2 uA	75.8 uA	23.7%	54.2%
FinFET	nSi pSiGe	1 GPa SRB	EPI	5	5.5	71 mV/dec	66.7 uA	71.5 uA	23.9%	52.8%
NWs 3	nSi pSiGe	0.5 GPa S/D stress	WPC	7	5.5	61 mV/dec	53 uA	60 uA	21%	56.1%
NWs 2	nSi pSiGe	0.5 GPa S/D stress	WPC	7	5.5	61 mV/dec	37 uA	42 uA	18%	58 %
FinFET	nSi pSi	0.5 GPa S/D stress	EPI	5	4.5	71 mV/dec	60.8 uA	63.4 uA	20%	56%

RO FO3 INV
Wire length 300CGP
NFIN= 3
Vdd= 0.65V
Rbeol = 135 ohm/um
Cbeol = 0.16 fF/um

LVT device, all devices at same Iddq 15nA
Lg=18nm, EOT=1.1nm, spacer=5nm
PP=42nm, MP=32nm
FH=35nm, FW=5nm

FIN HEIGHT, NUMBER OF STACKED WIRES

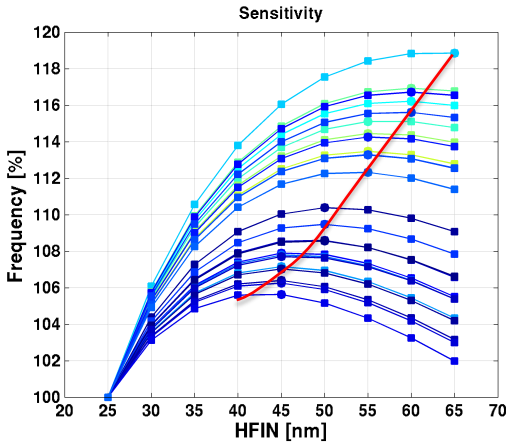
Increasing drive current by increasing effective width (Fin height, Number of wires) also increases

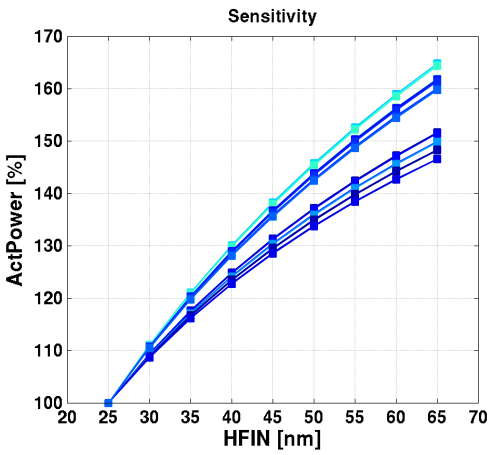
1) FEOL capacitances (gate, parasitics)
2) Voltage drop over Rseries (increase in drive current)
3) Dynamic power

From (1) and (2)

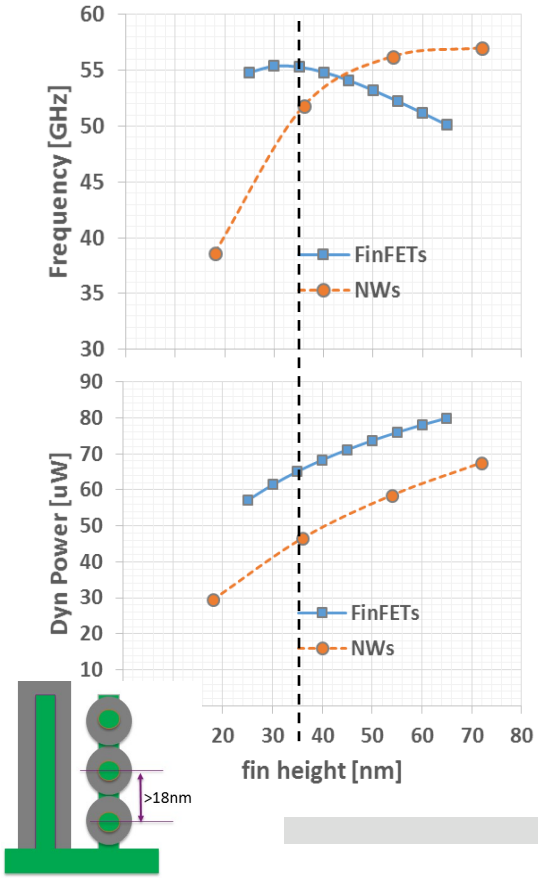
- the frequency increase saturates
- is always smaller than the dynamic power increase

FO3, 300CGP, for different Cfeol and Rseries



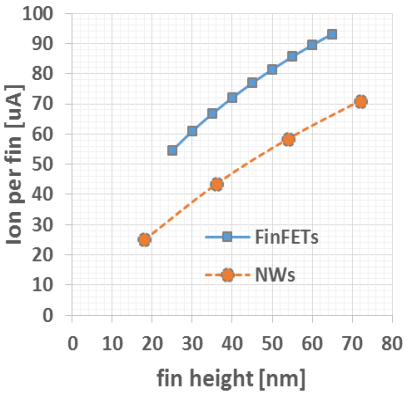


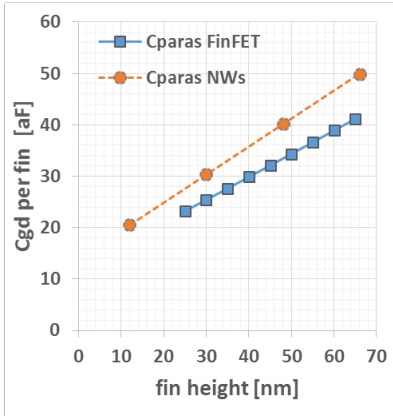
FINFET VERSUS NANOWIRES

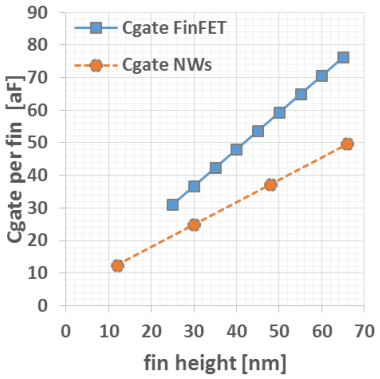


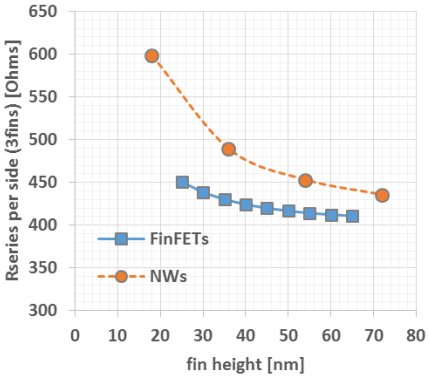
- Both Fin and NW performance saturate when increasing height and number of NW, respectively, while power keeps increasing.
  - Increased FEOL capacitance
  - Larger voltage drop over Rseries (because of higher current)
- FinFET vs. Nanowire
  - At same "height", finFET beats NW.
  - To beat FinFET need 3NW and taller structure with additional process complexity
  - NW offer a power advantage (even at same speed) due to lower gate capacitance.

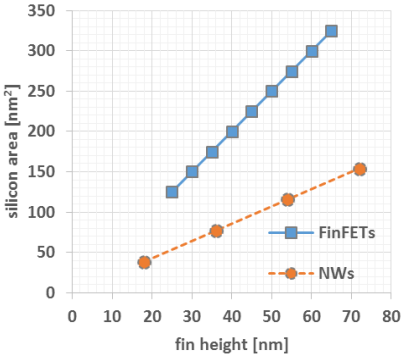
FINFETS VERSUS NANOWIRES











- Increase in current density at same Vdd due to SS (61 mV/dec vs 70 mV/dec)
- Benefit of quantization on charges / transport

- Penalty due to smaller silicon cross section vs. process complexity of stacking
- Penalty of increased series resistance (especially in extension)

DEVICE LEVEL CONCLUSIONS

Scaling to N7 has many paths, key requirements include:

- Stress Engineering
- Contact Engineering
- Gate spacer

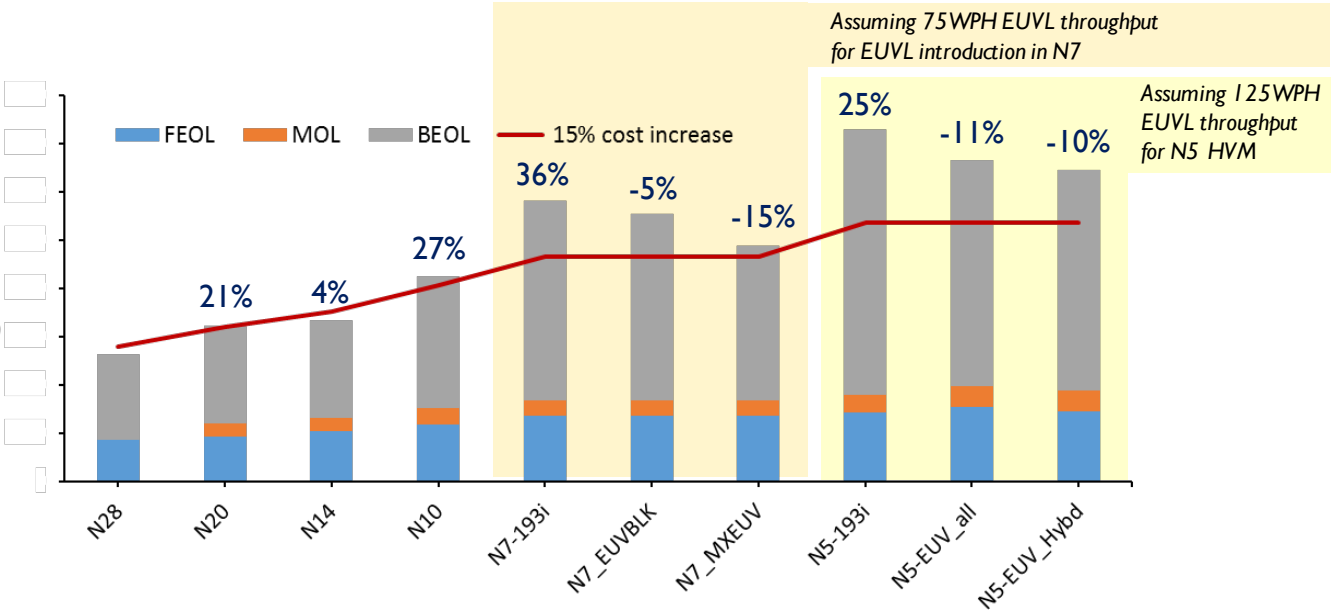
Fin vs NWs

- have similar performance as FinFET at N7 both in speed and power because of process complexity (NWs stacking) and saturation of performance with Fin height/ number of wires
- Will be more beneficial under smaller BEOL interconnect due to their smaller gate capacitance
- Would make large difference at N5 due to better electro-static behavior.

OVERALL IMPACT ON WAFER COST

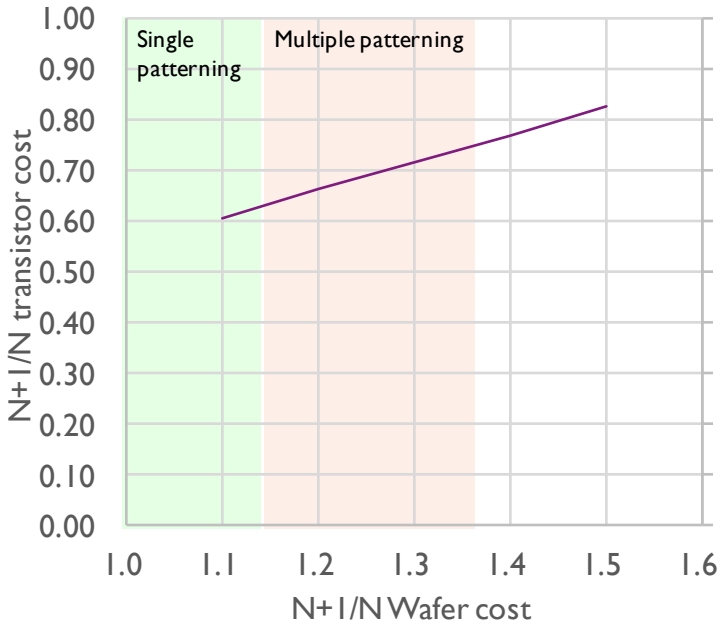
BEOL Stack:
M1, 3Mx, 6My, 2 WideMtl
Mint for N7-193i and N7 EUVBLK,
Mint for all N5 options.

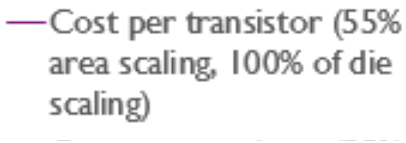
EUVL tool cost 1.6x of 193i tool, utilization of all litho tools assumed to be 60%, non-litho 70%; 193i TP = 200



- At wafer level, every patterning alternative costs significantly more than the 15% increase per node.
- N5 would need a matured EUVL technology with improved throughput
- A holistic scaling approach would help us to bring down the cost at a die level

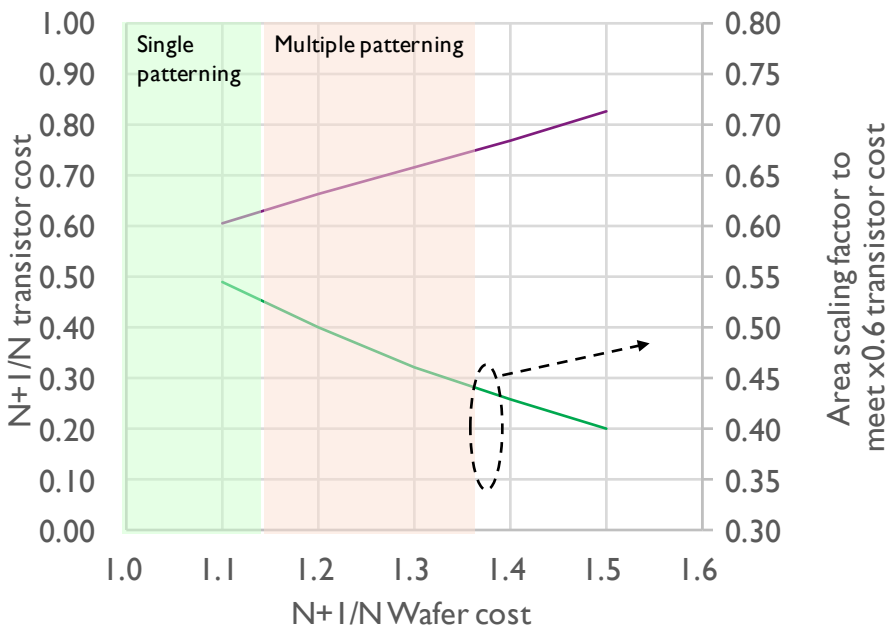
AREA SCALING IS CRITICAL TO STAY ON THE 0.6X COST/TRANSISTOR CURVE





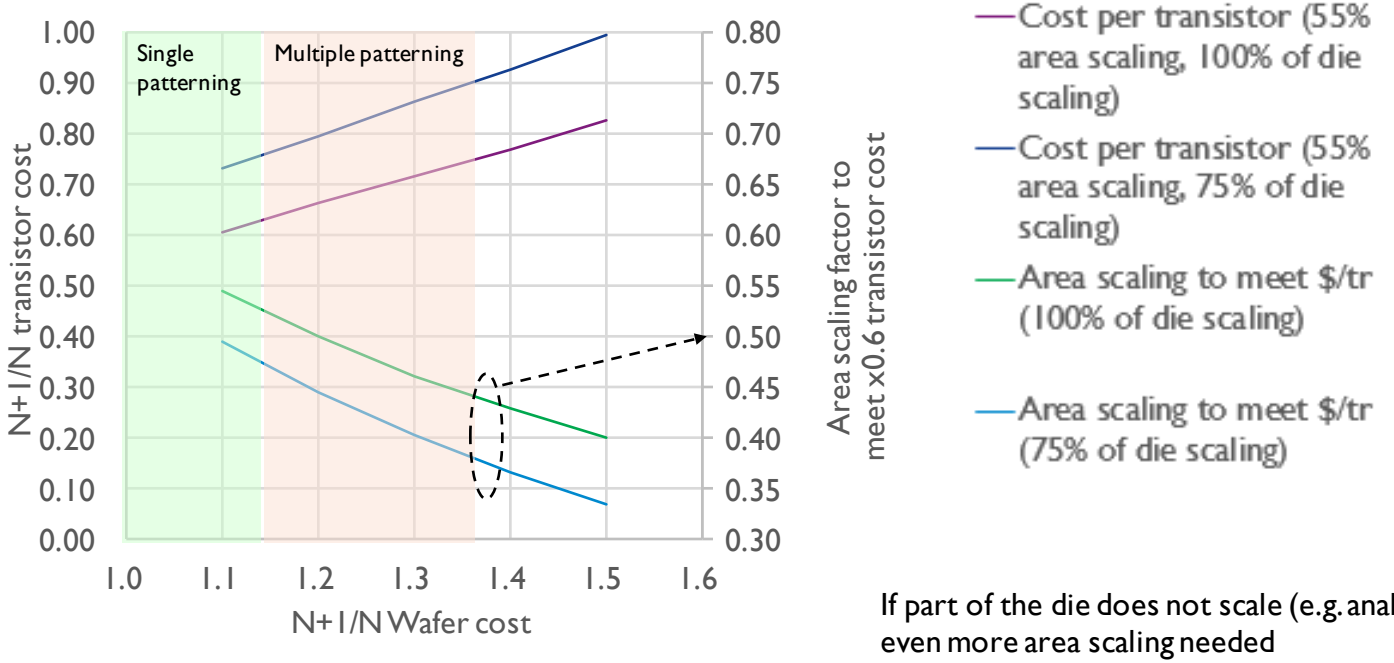
With multiple patterning, transistor cost scaling exceeds historical 0.6x target

AREA SCALING IS CRITICAL TO STAY ON THE 0.6X COST/TRANSISTOR CURVE



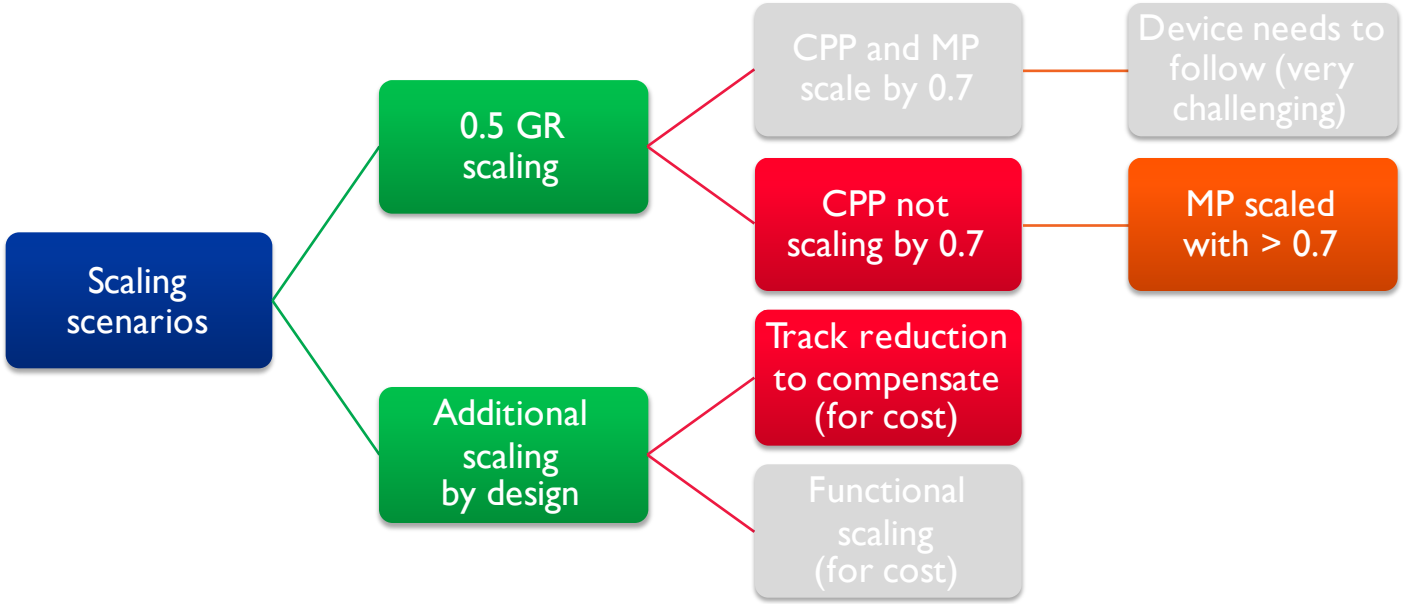
To compensate, area needs to scale with more than traditional x0.55

AREA SCALING IS CRITICAL TO STAY ON THE 0.6X COST/TRANSISTOR CURVE

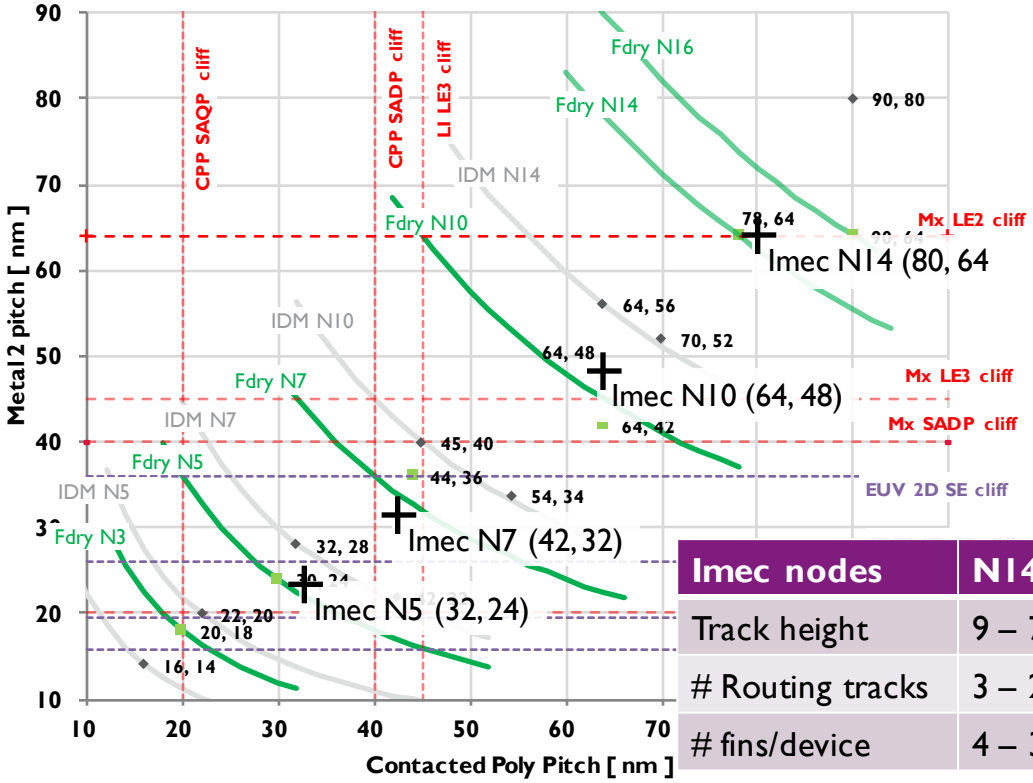


If part of the die does not scale (e.g. analog/RF), even more area scaling needed

LOGIC SCALING ROUTES



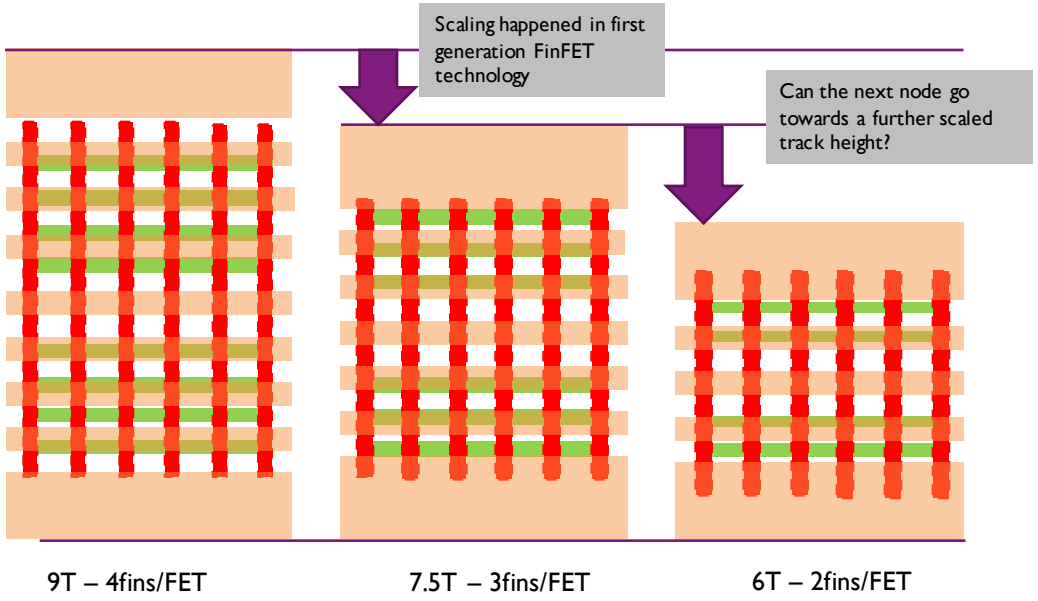
LOGIC SCALING LANDSCAPE



- MP x CPP determines logic cell area
- Choices of MP, CPP
  - Electrical performance
    - Reduced CPP forces trade-off between device electrostatics (gate-length) and parasitics (spacer, contact sizes)
    - Reduced MP leads to steep increase in BEOL resistance
  - Patterning cliffs, cost

Imec nodes	N14	N10	N7	N5
Track height	9 – 7.5	7.5	7.5	7.5 – 6
# Routing tracks	3 – 2	4	3-4	3- 2
# fins/device	4 – 3	3 - 2	3 - 2	2- 1

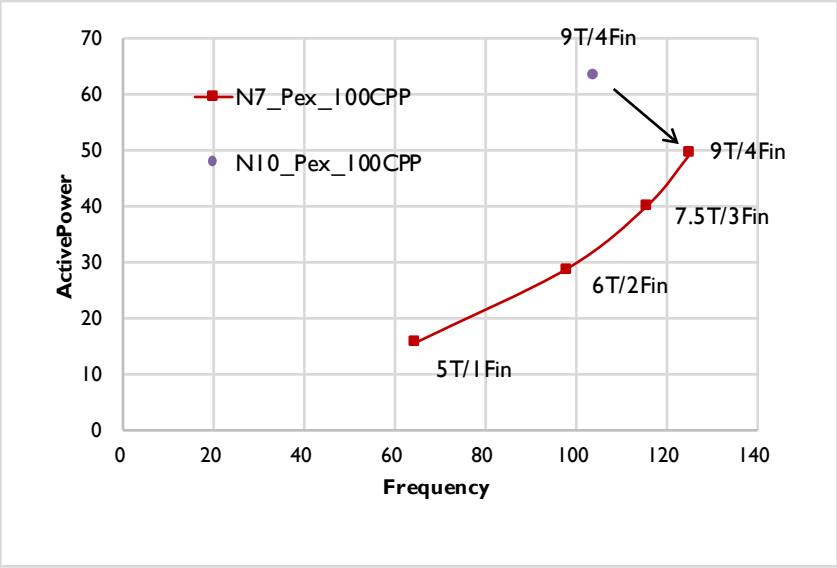
TRACK SCALING TO GET EXTRA SCALING



CPP	MP	FP	SDC Tracks	SDC fins
64	48	36	7.5T	3
42	32	24	7.5T	3
42	32	24	6T	2
x0.65	x0.66	x0.66	0.43 (7.5T) 0.34 (6T)	

- 7.5T cells
  - traditional scaling path
- 6T cells
  - more aggressive area scaling
  - Performance depends on reduction of device parasitics to benefit from taller fin
  - Good for power.

TRACK SCALING TO GET EXTRA SCALING



RO FO3 INV
Wire length 100CGP
Vdd= 0.65V
Rbeol = 135 ohm/um
Cbeol = 0.16 fF/um

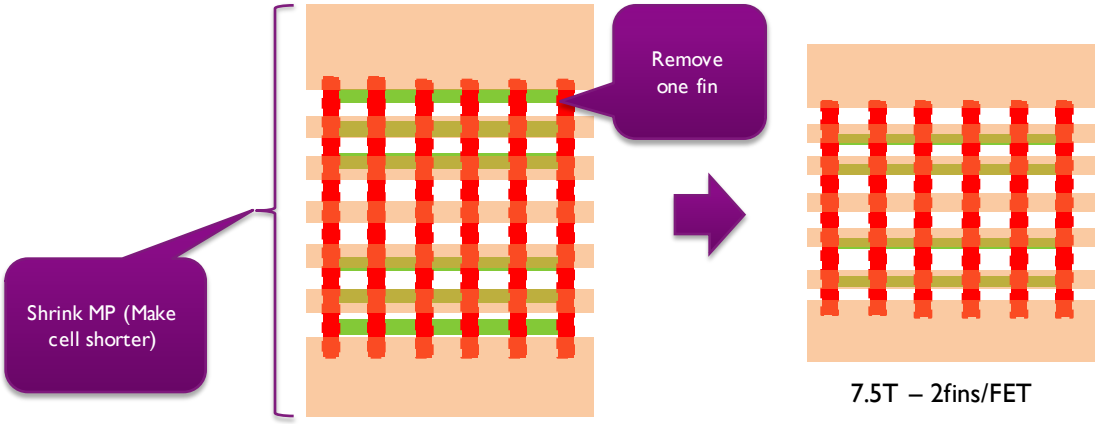
LVT device, all devices at same Iddq
Lg=18nm, EOT=1.1nm, spacer=6nm
PP=42nm, MP=32nm
FH=35nm
FW=5nm

- Keeping 7.5 tracks while scaling pitches
  - Traditional scaling path
  - Area scaling with x0.43
  - Performance and power benefit
- Reducing track height to 6 tracks with respect to previous node (7.5 tracks)
  - gives extra area scaling at cell level (~15-20%)
  - Reduce power
  - However at the cost of limited/no performance gain (wrt previous node)

AGGRESSIVE METAL PITCH SCALING

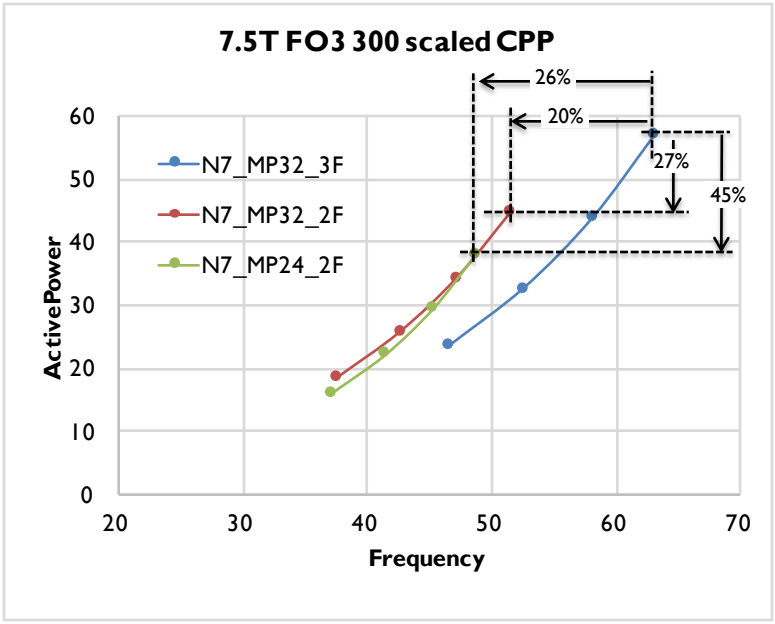
- Push SAQP on Mx to the limit
- However don't scale FP as aggressive to keep gate process simpler
- Scaling of Gate Pitch limited

CPP	MP	FP	SDC Tracks	SDC fins
64	48	36	7.5T	3
42	24	24	7.5T	2
x0.65	x0.5	x0.66	x0.33	



- Aggressive area scaling x0.33
- Performance would depend on ability to reduce device parasitics to benefit from taller fins.
- Good for power.
- Status quo for PnR

AGGRESSIVE METAL PITCH SCALING



RO FO3 INV
Vdd= 0.65V

LVT device, all devices at same lddq
Lg=18nm, EOT=1.1nm, spacer=6nm
PP=42nm, FH=35nm ,FW=5nm

- Aggressive MP scaling would allow to obtain aggressive power savings node to node of around 70%
  - 45% from N7 MP32 to N7 MP24
    on top of ~45% savings from N10 to N7
- However at the cost of limited/no performance gain (wrt previous node)
  - 26% drop in performance from N7 MP32 to N7 MP24 is about
    equal to performance gain from N10 to N7
- Performance gap can be further improved by continued contact and mobility enhancement and BEOL Cap reduction

CONCLUSIONS

Going to heavy multi-patterning era needs 'extra' scaling

- But multi-patterning has a lot of technology development
- To compensate for this need for Design-Technology-Device cooptimization
- Implies more technology half/quarter nodes/(dead nodes)

Scaling beyond 0.5 needed to keep Moore's law alive

- However we can-not gain in power "and" performance "and" cost "and" area
- Technologies likely to diversify to give more gains in one or other metric