THE CHIPS TO SYSTEMS CONFERENCE

SHAPING THE NEXT GENERATION OF ELECTRONICS

JUNE 23-27, 2024

MOSCONE WEST CENTER
SAN FRANCISCO, CA, USA

THE CHIPS TO SYSTEMS CONFERENCE
SHAPING THE NEXT GENERATION OF ELECTRONICS

JUNE 23-27, 2024
MOSCONE WEST CENTER
SAN FRANCISCO, CA, USA

SIEMENS

Lesyn: Placement-Aware Logic Resynthesis for Non-Integer Multiple-Cell-Height Designs

Yuan Pu[1], Fangzhou Liu[1], Yu Zhang[1], Zhuolun He[1], Yibo Lin[2], Kai-Yuan Chao[3], Bei Yu[1]

[1]Chinese University of Hong Kong
[2]Peking University
[3]Siemens

Outline

Introduction

Introduction

- Standard cell libraries → varying cell heights → diverse PPA requirements.
- Traditional multi-row-height cells → sized as integer multiples of a single-row height → two drawbacks[1]:
  - Unnecessary area cost.
  - Cells not align with the shrinking transistor sizes and minimal wiring pitch.
- Non-integer Multi-Cell Height (NIMCH)[2] technology provides:
  - More extensive design space for exploration.
  - Highly flexible and efficient power, performance, and area co-optimization.
  - 8T cells (shorter) and 12T cells (higher)[3].

[1] M. Hatamian and P. Penzes, "Non-integer height standard cell library", US Patent 8,788,998.

[2] S.-Y. Wu *et al.*, "A 3nm CMOS FinFlex™ Platform Technology with Enhanced Power Efficiency and Performance for Mobile SoC and High Performance Computing Applications", in *Proc. IEDM*, IEEE, 2022, pp. 27–5.

[3] S. A. Dobre *et al.*, "Design implementation with noninteger multiple-height cells for improved design quality in advanced nodes", *IEEE TCAD*, vol. 37, no. 4, pp. 855–868, 2017.

NIMCH Placement

- New layout constraint:
  - Standard cells with the same cell heights are constrained to be located in the same sub-region.
  - Two categories based on the sub-region pattern: row-based and island-based. based on their cell heights

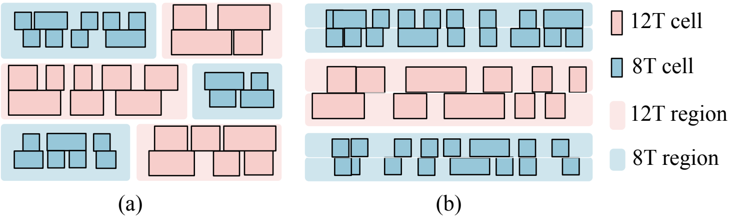
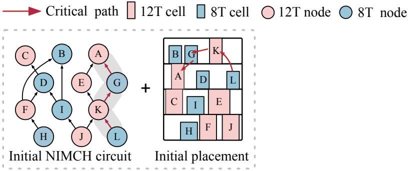


Illustration of (a) island-based and (b) row-based NIMCH sub-region patterns.

NIMCH Placement: Previous Work[4,5,6]

Conventional NIMCH physical design flow:

- Step 1: Initial placement (8T/12T cells spatially mixed).
- Step 2: NIMCH placer: form separate 8T/12T sub-regions



Drawbacks: Huge displacement of cells and timing degradation.

[4] S. A. Dobre *et al.*, "Design implementation with noninteger multiple-height cells for improved design quality in advanced nodes", *IEEE TCAD*, vol. 37, no. 4, pp. 855–868, 2017.

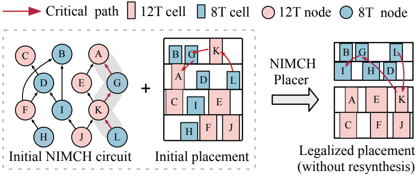
[5] Z.-Y. Lin and Y.-W. Chang, "A row-based algorithm for non-integer multiple-cell-height placement", in *Proc. ICCAD*, IEEE, 2021, pp. 1–6.

[6] J. Chen *et al.*, "An efficient epist algorithm for global placement with non-integer multiple-height cells", in *Proc. DAC*, IEEE, 2020, pp. 1–6.

NIMCH Placement: Previous Work[4,5,6]

Conventional NIMCH physical design flow:

- Step 1: Initial placement (8T/12T cells spatially mixed).
- Step 2: NIMCH placer: form separate 8T/12T sub-regions



Drawbacks: Huge displacement of cells and timing degradation.

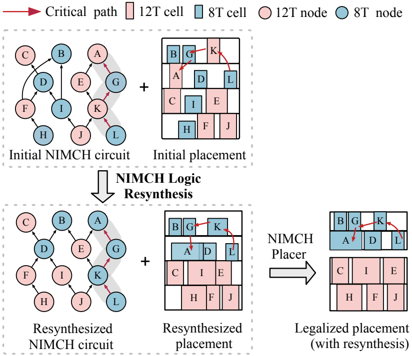
[4] S. A. Dobre *et al.*, "Design implementation with noninteger multiple-height cells for improved design quality in advanced nodes", *IEEE TCAD*, vol. 37, no. 4, pp. 855–868, 2017.

[5] Z.-Y. Lin and Y.-W. Chang, "A row-based algorithm for non-integer multiple-cell-height placement", in *Proc. ICCAD*, IEEE, 2021, pp. 1–6.

[6] J. Chen *et al.*, "An efficient epist algorithm for global placement with non-integer multiple-height cells", in *Proc. DAC*, IEEE, 2020, pp. 1–6.

NIMCH Placement: Motivation

- Intuition: Maintain the positions of cells in the initial placement, and modify the cell type/height.
- Logic resynthesis: Locally modify cell type/height, while preserving functionality.
- Potential solution: Insert logic resynthesis into the conventional NIMCH flow → between initial placer and NIMCH placer.



Problem Formulation

Row-Based NIMCH Logic Resynthesis

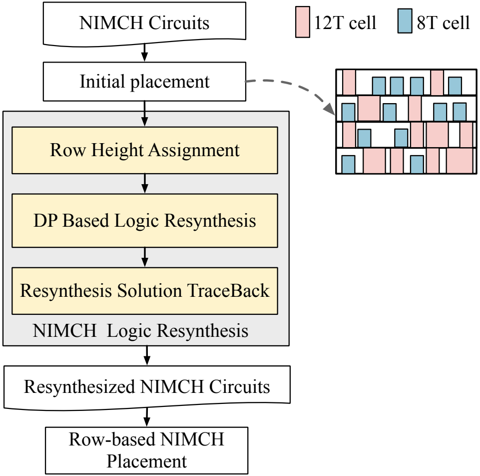
Given a circuit C with n gates, a set of library gates L and the initial placement of C, resynthesize each gate in C to minimize the combined cost of total cell area and maximal delay, and ensure the placement of the resynthesized circuit aligns with the row-based sub-region pattern.

- Reasons for area minimization:
  - Large cell area $\rightarrow$ cell inflation on placement rows $\rightarrow$ displacement for NIMCH legalization.
  - Large cell area $\rightarrow$ large utilization rate $\rightarrow$ routing issues.

Algorithm

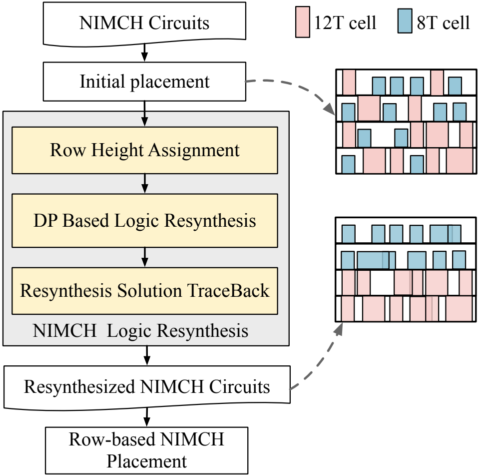
Overall Flow

- Three stages:
  - Row height assignment: a specific cell height assigned to each cell → the row-based sub-region pattern.
  - DP Based Logic Resynthesis.
  - Resynthesis Solution Traceback.
- Two categories of logic resynthesis:
  - Conservative resynthesis: The gate type is modified while the circuit topology remains unchanged.
  - Structural resynthesis: Both gate type and circuit topology are modified.



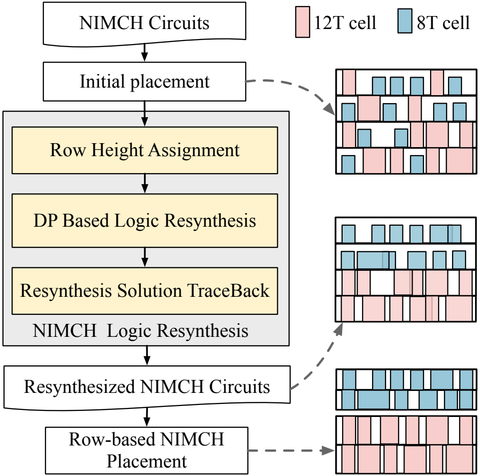
Overall Flow

- Three stages:
  - Row height assignment: a specific cell height assigned to each cell → the row-based sub-region pattern.
  - DP Based Logic Resynthesis.
  - Resynthesis Solution Traceback.
- Two categories of logic resynthesis:
  - Conservative resynthesis: The gate type is modified while the circuit topology remains unchanged.
  - Structural resynthesis: Both gate type and circuit topology are modified.



Overall Flow

- Three stages:
  - Row height assignment: a specific cell height assigned to each cell → the row-based sub-region pattern.
  - DP Based Logic Resynthesis.
  - Resynthesis Solution Traceback.
- Two categories of logic resynthesis:
  - Conservative resynthesis: The gate type is modified while the circuit topology remains unchanged.
  - Structural resynthesis: Both gate type and circuit topology are modified.



DP-Based Logic Resynthesis

- Objective: Minimize the total cell area.
- Constraints:
  - The maximal delay between any primary input and primary output is less than the target delay α;
  - The cell height of each gate g_i satisfies the row height assignment solution RHA.
- Two steps:
  - DP Entry Calculation.
  - Solution traceback.

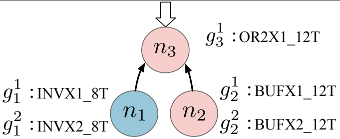
$$\min \quad \sum_{i=1}^{N} A_i \quad \text{s.t.} \quad \text{Delay} \le \alpha, \quad g_i \in L, \quad h(i) = RHA[i], \tag{1}$$

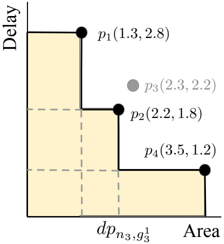
Deinition: Fanin Mapping & AD Curve

Definition

(fanin mapping) For a node n in a circuit, denote its fanin nodes by $Fanins(n)$ and for each fanin node k ($k \in Fanins(n)$), $Gates(k)$ denotes a set of library gates that can be mapped to k. A ***fanin mapping*** *for node n, denoted by p_n, refers to a library gate mapping assignment for all fanin nodes of n ($p_n = \{k \to g_k, \forall k \in Fanins(n), g_k \in Gates(k)\}$).*

$$p_1 = \{n_1 \to g_1^1, n_2 \to g_2^1\} \to \{area = 1.3, delay = 2.8\}$$
$$p_2 = \{n_1 \to g_1^1, n_2 \to g_2^2\} \to \{area = 2.2, delay = 1.8\}$$
$$p_3 = \{n_1 \to g_1^2, n_2 \to g_2^1\} \to \{area = 2.3, delay = 2.2\}$$
$$p_4 = \{n_1 \to g_1^2, n_2 \to g_2^2\} \to \{area = 3.5, delay = 1.2\}$$





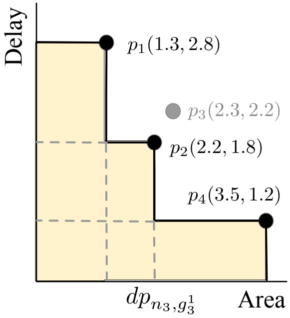
(a) Illustration of fanin mappings of n_3 (b) AD-curve of n_3 and g_3^1

DP Entry Calculation

For any node n_i in the circuit with library gate g_i and fanin mapping p_i, to determine the *delay* and *area* values of dp_{n_i,g_i,p_i}.

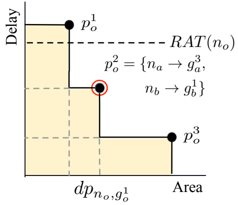
$$dp^{(\text{delay})}_{n_i,g_i,p_i} = D(g_i) + \max_{n_f,g_f} d^{(i)}_{\min}[(n_f,g_f)],$$
$$dp^{(\text{area})}_{n_i,g_i,p_i} = \frac{A(g_i) + \sum_{g_f \in p_i} A(g_f)}{1 + |\text{Fanins}(n_i)|}, \tag{2}$$

AD-curve construction.



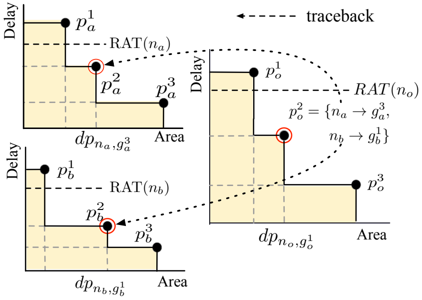
Solution Traceback

- Phase: After dp-entry calculation.
- Reverse-topological order, and starts with the nodes whose fanouts are POs (output nodes for short).
- Merge all AD-curves of dp_{n_o,g_o^i} ($g_o^i \in G_0$) into a single AD-curve, and among the points whose delays are smaller than the required arrival time of n_o (RAT(n_o)), the point with the minimal *area*, is selected as the resynthesized solution of n_o.



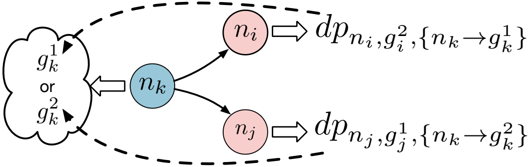
Solution Traceback

- Phase: After dp-entry calculation.
- Reverse-topological order, and starts with the nodes whose fanouts are POs (output nodes for short).
- Merge all AD-curves of dp_{n_o,g_o^i} ($g_o^i \in G_0$) into a single AD-curve, and among the points whose delays are smaller than the required arrival time of n_o (RAT(n_o)), the point with the minimal *area*, is selected as the resynthesized solution of n_o.



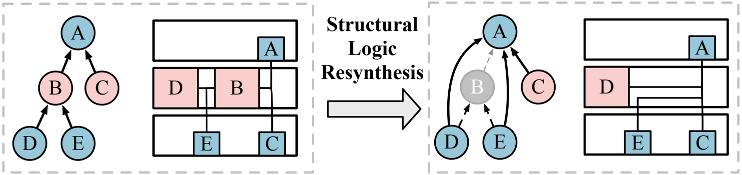
Solution Traceback: Path Reconvergence

- When two nodes (n_i and n_j) share the same fanin node (n_k), and the dp table entry determination of n_i and n_j leads to gate-type-selection conflict of n_k.
- Solution: if the gate-selection conflict occurs for node n_k, we calculate the required arrival time of n_k (RAT(n_k)), and between the two candidate library gates, we choose the candidate with the minimal area whose *delay* is smaller than RAT(n_k).



Structural Resynthesis

- Allow circuit topological modification during logic resynthesis.
- For larger solution space and potentially better resynthesis result.

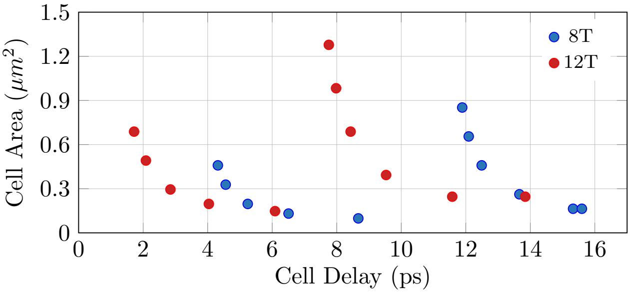


(a) (b) (c) (d)

Experimental Results

Experimental Setting

- Benchmark: EPFL combinational arithematic suite.
- PDK: Modified 15nm FinFET-based Open Cell Library (OCL).
- Initial placement: Innovus.



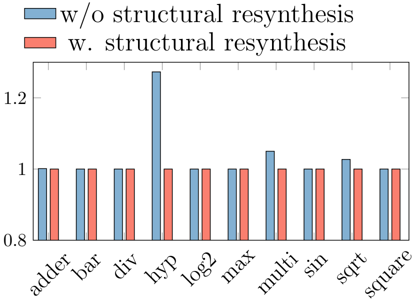
Experimental Results

Table: Post-Place PPA results on the EPFL combinational arithmetic benchmark suite. 'Conventional flow' involves directly applying the row-based NIMCH placer on the initial placement. For 'conventional flow +Lesyn', our proposed resynthesis approach is incorporated between the stage of initial placement generation and row-based NIMCH placement.

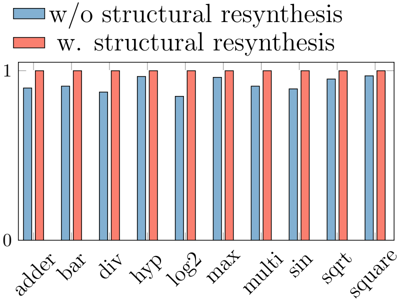
Circuit	Metrics		conventional flow [7]					conventional flow + Lesyn (Ours)				
	# Cells	# Nets	area (um^2)	power (mW)	delay (ns)	disp (um)	time (s)	area (um^2)	power (mW)	delay (ns)	disp (um)	time (s)
adder	891	1147	197.198	0.012	1.865	2410.82	47	197.231	0.011	1.715	923.65	59
bar	2188	2323	744.096	0.043	1.525	4281.73	54	727.213	0.039	0.485	311.30	77
div	48343	48471	9930.834	0.549	245.849	185512.02	1248	9850.830	0.525	208.065	6223.10	2182
hyp	168035	168291	38847.320	2.181	1154.000	4623408.61	3943	38923.400	2.082	795.058	30046.210	6505
log2	21148	21180	5638.914	0.334	21.330	102082.70	704	5546.227	0.324	12.821	5296.64	1307
max	2287	2799	493.994	0.025	2.839	2267.18	83	490.160	0.024	2.652	515.07	103
multi	21832	21960	4415.000	0.248	9.996	395585.00	493	4443.712	0.246	5.646	2186.75	836
sin	4203	4227	1212.678	0.072	3.710	13039.92	165	1219.117	0.072	3.269	540.42	252
sqrt	21832	21960	5962.334	0.355	147.273	168311.07	870	5942.411	0.361	142.552	2803.20	1377
square	12985	13050	2705.146	0.135	6.155	35430.74	240	2704.163	0.134	5.748	1944.83	361
Normalize	-	-	1.000	1.000	1.000	1.000	**1.000**	**0.999**	**0.965**	**0.739**	**0.009**	1.538

[7] Z.-Y. Lin and Y.-W. Chang, "A row-based algorithm for non-integer multiple-cell-height placement", in *Proc. ICCAD*, IEEE, 2021, pp. 1–6.

Analysis of Structural Resynthesis



(c) maximal delay comparison



(d) runtime comparison

Comparison of normalized maximal delay and runtime, with and without structural logic resynthesis.

THE CHIPS TO SYSTEMS CONFERENCE

SHAPING THE NEXT GENERATION OF ELECTRONICS

JUNE 23-27, 2024

MOSCONE WEST CENTER
SAN FRANCISCO, CA, USA

SIEMENS

THANK YOU!