Bounded Potential Slack: Enabling Time Budgeting for Dual-V_t Allocation of Hierarchical Design

Jun Seomun, Seungwhun Paik
and Youngsoo Shin
Dept. of Electrical Engineering, KAIST, KOREA

Contents

- Introduction
  - Time budgeting
  - Motivation
- BPS-based time budgeting
  - Bounded potential slack
  - LP formulation
  - Design flow
- Experimental results
- Conclusion

Hierarchical VLSI Design

- Hierarchical design
  - Maintain hierarchy through RTL design down to layout
  - Popular in μ-processor designs and SoC designs

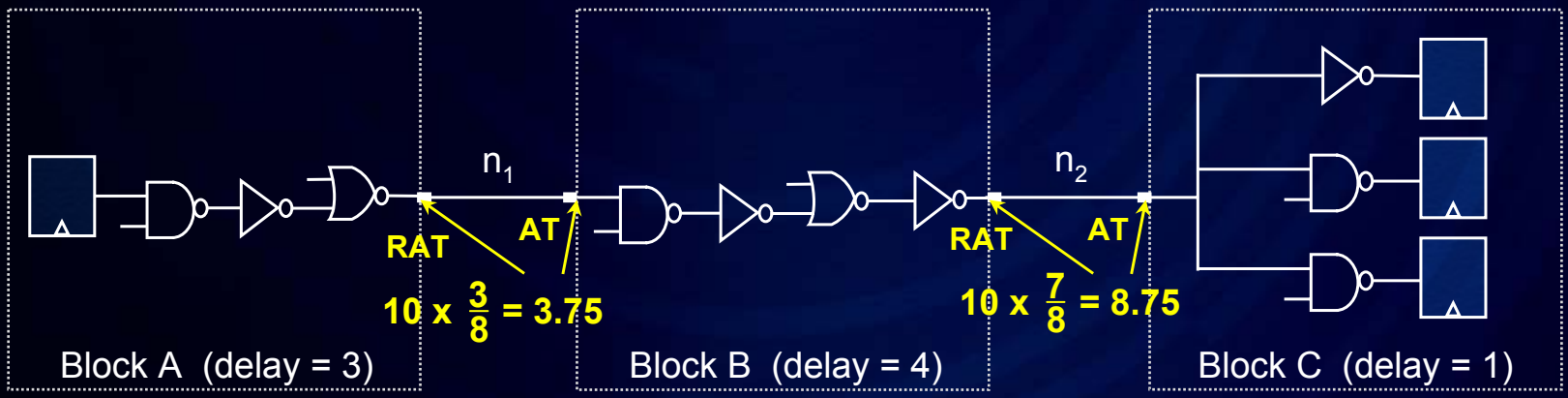
An example hierarchical design: Alpha processor

- Advantages
  - Facilitate parallel development
  - Facilitate IP reuse
  - Less complexity of physical design

- Disadvantages
  - More design steps (floorplanning, pin assignment, wire resource allocation, time budgeting)

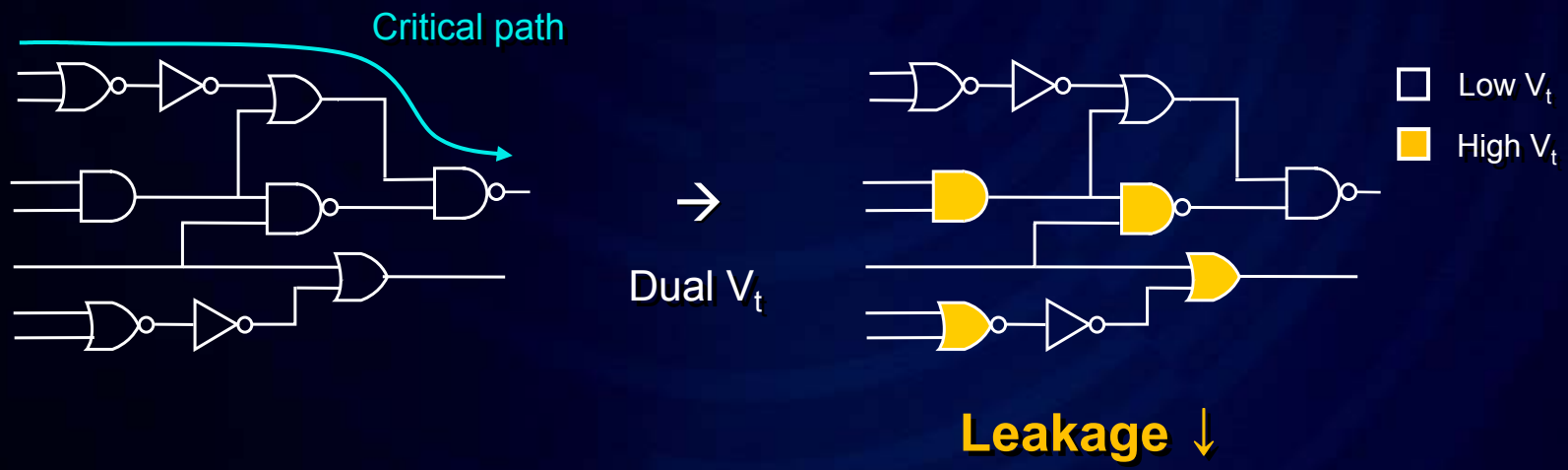
Time Budgeting

- Generate timing assertions at each block boundary
  - Arrival time/slew limit @ block input
  - Required arrival time/load limit @ block output
- Conventional time budgeting
  - Distribute slack to each block in proportion to its own delay
  - Example
    - Clock period : 10



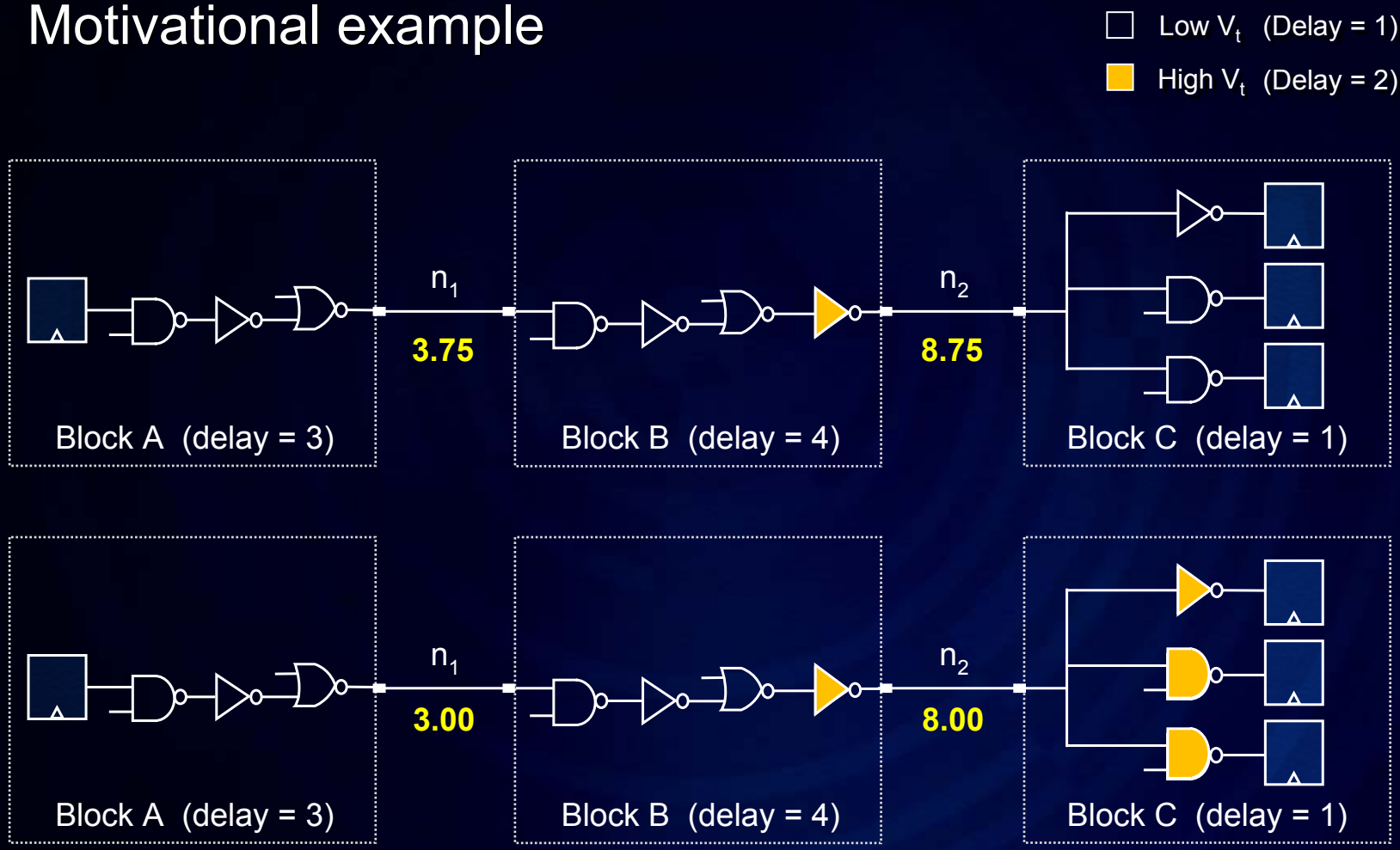
Dual V_t

- Dual V_t
  - Can be integrated into conventional design flow
  - Reduce both active and standby leakage
- Dual V_t allocation
  - Critical path →Low V_t (fast, large leakage)
  - Non-critical → High V_t (slow, small leakage)



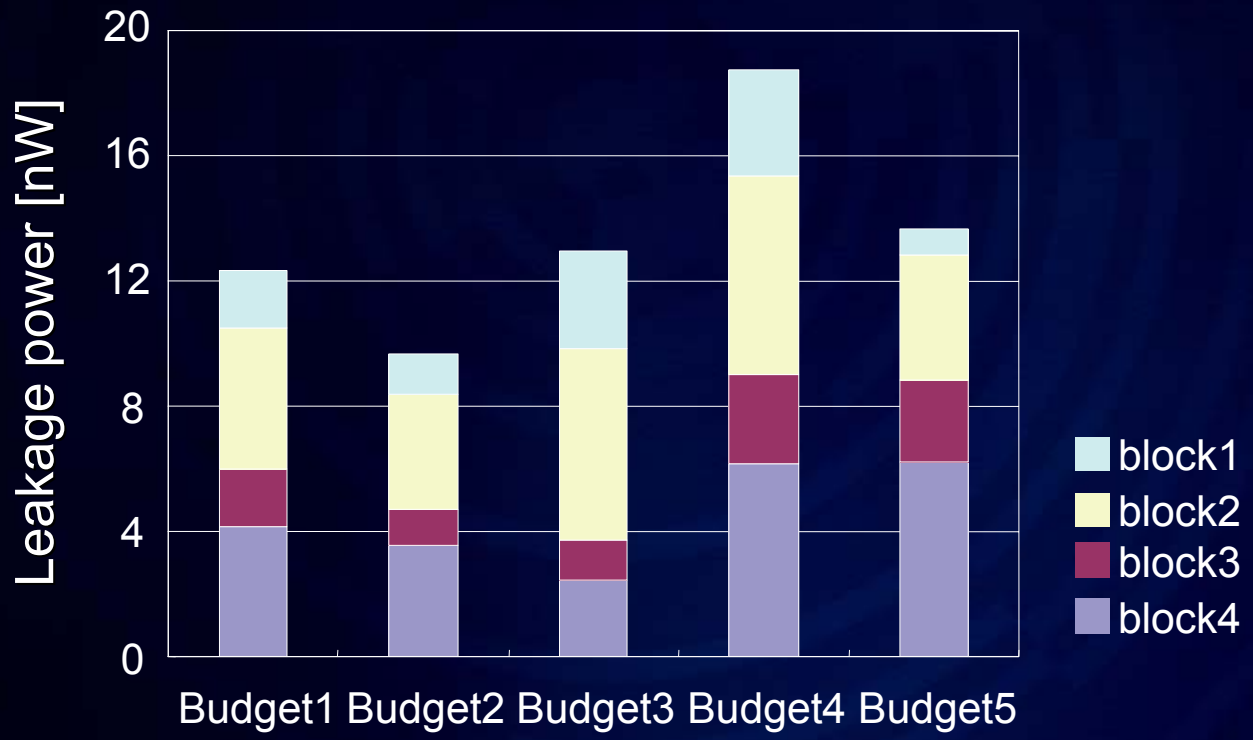
Effect of Time Budgeting on Dual V_t

- Motivational example



Effect of Time Budgeting on Dual V_t

- Proportion of high V_t gates is dictated by timing assertion



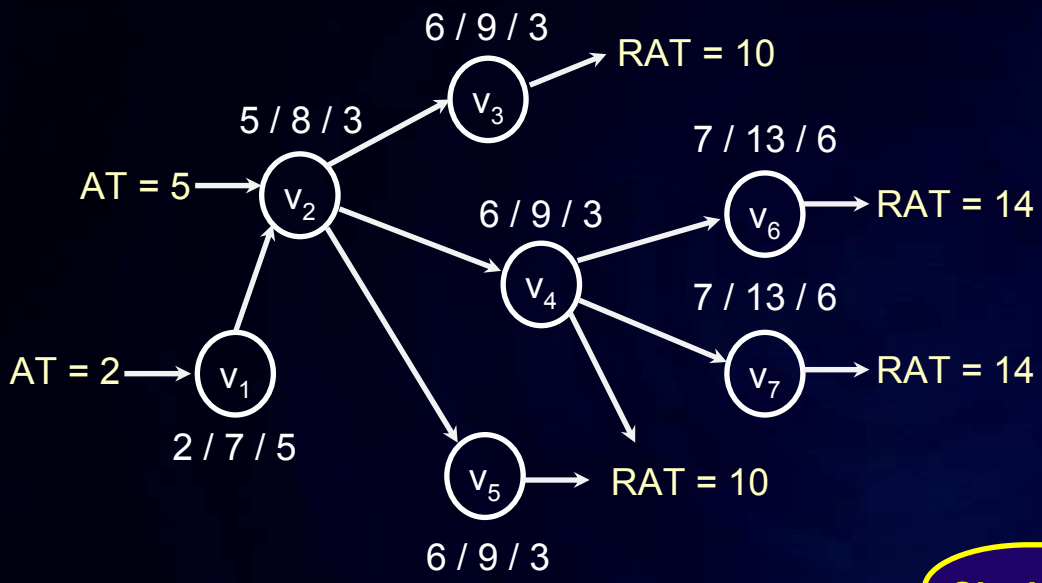
Key Idea

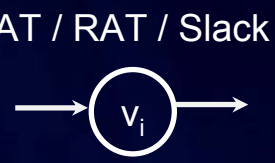
- Bounded potential slack (BPS)
  - Potential slack under the limited incremental delays
  - Strongly correlated with the proportion of high V_t gates
  - Use as a measure of dual V_t allocation

- Dual V_t aware time budgeting
  - BPS-based time budgeting
  - Achieving maximum bounded potential slack
  - Linear programming

Potential Slack & Bounded Potential Slack

- Notation





Delay vector $D = [\, d_1 \quad d_2 \quad \dots \quad d_n \,]$

Slack vector $S = [\, s_1 \quad s_2 \quad \dots \quad s_n \,]$

Incremental delay vector $\Delta D = [\, \Delta d_1 \; \Delta d_2 \dots \Delta d_n \,]$

$D = [\, 1\ 1\ 1\ 1\ 1\ 1\ 1 \,]$

$S = [\, 5\ 3\ 3\ 3\ 3\ 6\ 6 \,]$

Slack assignment

Deriving $\Delta D = [\Delta d_1 \; \Delta d_2 \dots \Delta d_n \,]$

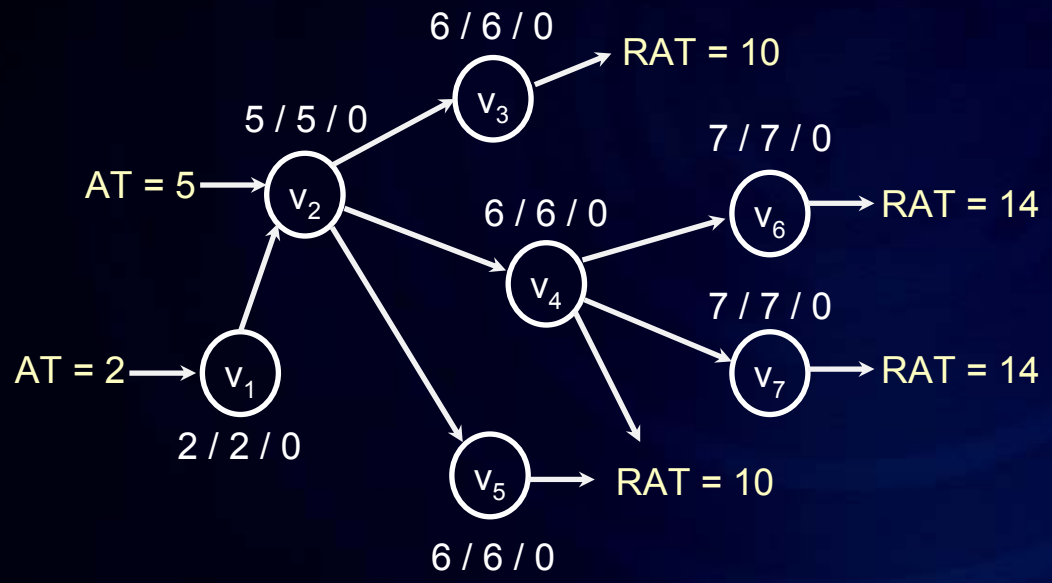
$D = [d_1 + \Delta d_1 \dots d_n + \Delta d_n \,]$

* C. Chen et al., "Potential Slack: an Effective Metric of Combinational Circuit Performance", ICCAD 2000

Potential Slack & Bounded Potential Slack

- Potential slack
  - Maximum value of the sum of incremental delays

$$\text{Max.} \sum_{i=1}^{n} \Delta d_i$$



$\Delta D = [2\ 0\ 3\ 0\ 3\ 6\ 6]$

$D = D + \Delta D = [3\ 1\ 4\ 1\ 4\ 7\ 7]$

$S = [0\ 0\ 0\ 0\ 0\ 0\ 0]$

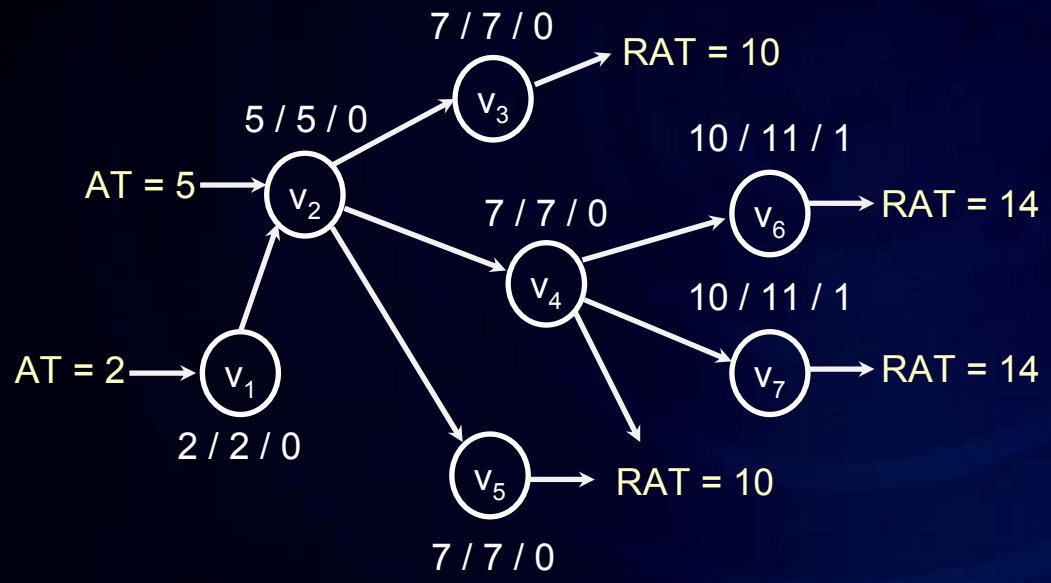
Potential slack = 20

* C. Chen et al., "Potential Slack: an Effective Metric of Combinational Circuit Performance", ICCAD 2000

Potential Slack & Bounded Potential Slack

- Bounded potential slack
  - Maximum value of the sum of (bounded) incremental delays

$$\text{Max.} \sum_{i=1}^{n} \Delta d_i \quad where \quad 0 \le \Delta d_i \le B_i \quad , \quad B_i = d_i \big|_{V_t=high} - d_i \big|_{V_t=low}$$



Assume $B_i = 2$ for all gates

$\Delta D = [\ 2\ 1\ 2\ 2\ 2\ 2\ 2\ ]$

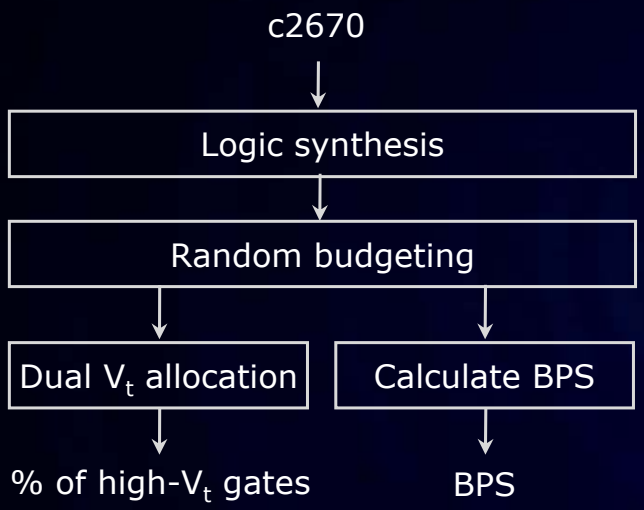
$D = D + \Delta D = [\ 3\ 2\ 3\ 3\ 3\ 3\ 3\ ]$

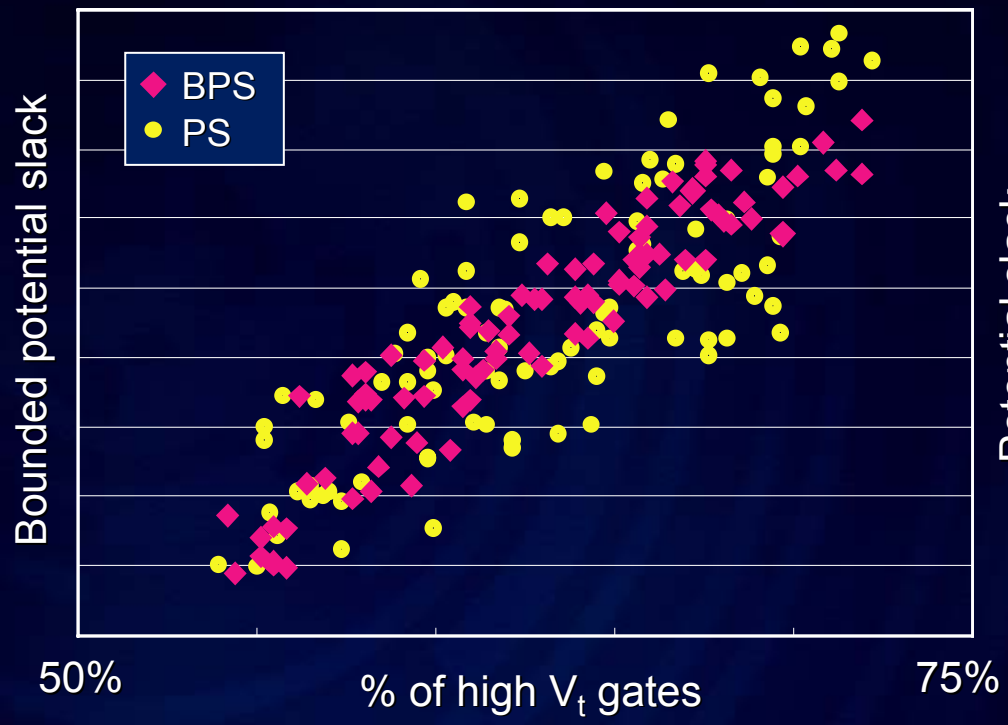
$S = [\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ ]$

Bounded potential slack = 13

Potential Slack & Bounded Potential Slack

- Correlation test





  - Strong correlation with the proportion of high-V_t gates
    - Correlation coefficient $\rho = 0.94$

BPS-Based Time Budgeting

- LP formulation
  - Objective function
    - Maximize bounded potential slack

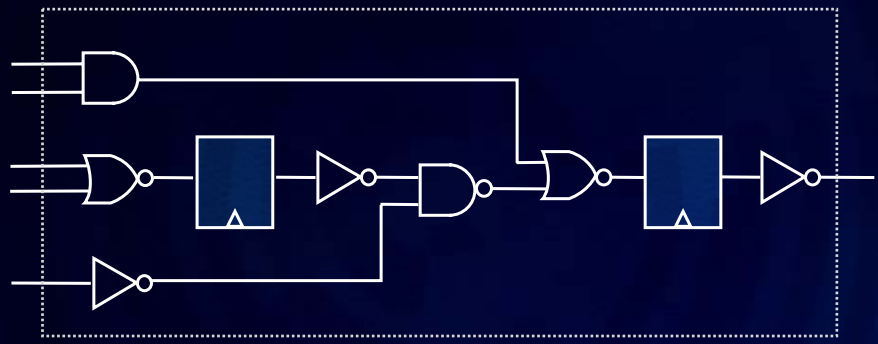
$$Maximize \sum_{n \in N} \sum_{i \in V_n} \Delta d_i$$

N : *Set of blocks*
V_n : *Set of gates in block n*

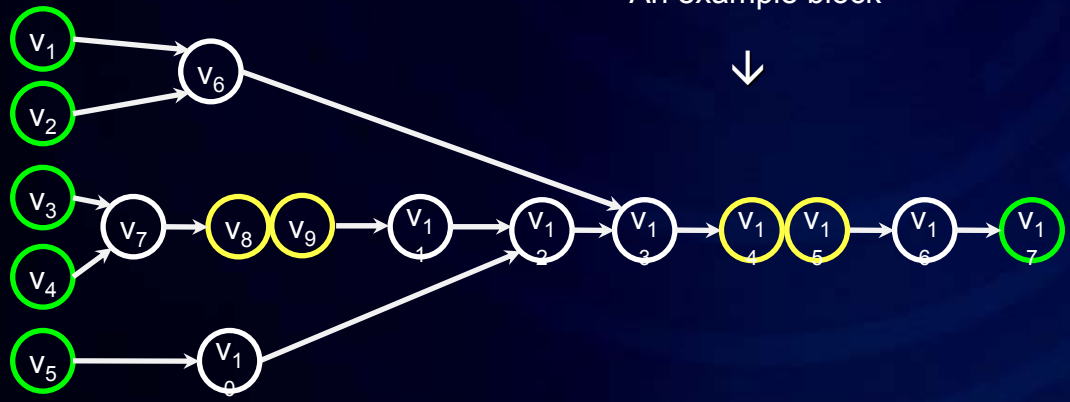
  - Constraints
    - Block boundaries
    - Flip-flop boundaries
    - Combinational gates
    - Bounds for incremental delay

BPS-Based Time Budgeting

- LP formulation
  - Graph modeling of a hierarchical block



An example block



BI (Block input)
BO (Block output)

FI (Flip-flop input)
FO (Flip-flop output)

CG (Combinational gate)

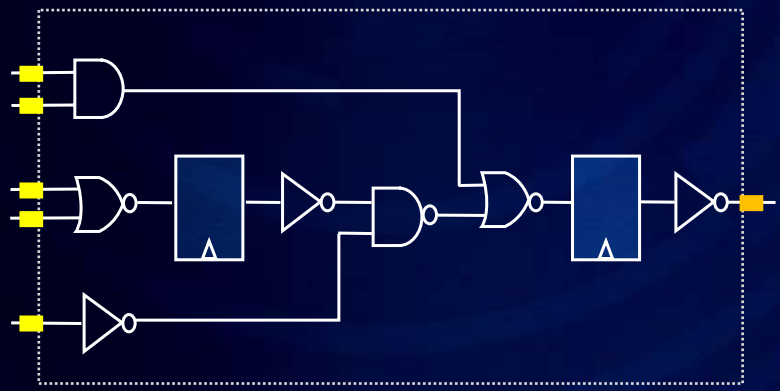
BPS-Based Time Budgeting

- LP formulation
  - Constraints for block boundaries
    - Block input arrival time

$$a_i \geq x_i \quad , \quad \forall v_i \in BI$$

    - Block output required arrival time

$$r_i \leq x_i \quad , \quad \forall v_i \in BO$$



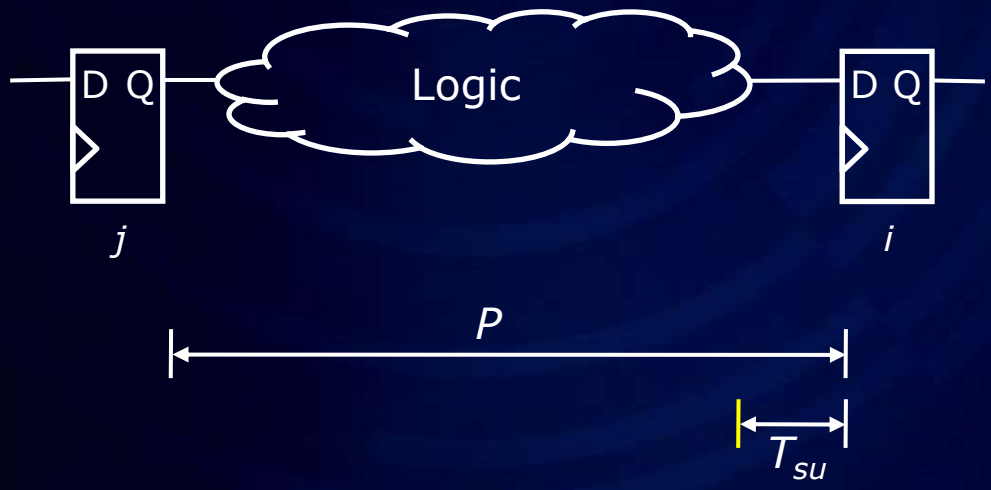
BPS-Based Time Budgeting

- LP formulation
  - Constraints for flip-flop boundaries
    - Flip-flop input required arrival time

$$r_i \leq P - T_{su} - \Delta d_i, \quad \forall v_i \in FI$$

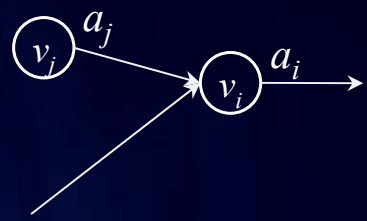
    - Flip-flop output arrival time

$$a_i \geq T_{cq} + \Delta d_i, \quad \forall v_i \in FO$$



BPS-Based Time Budgeting

- LP formulation
  - Constraints for combinational gates
    - Gate output arrival time



$$a_i \geq a_j + d_i + \Delta d_i, \quad \forall v_i \in CG$$

where v_j *is a fanin of* v_i

$$r_i - a_i \geq 0, \quad \forall v_i \in CG + BO$$

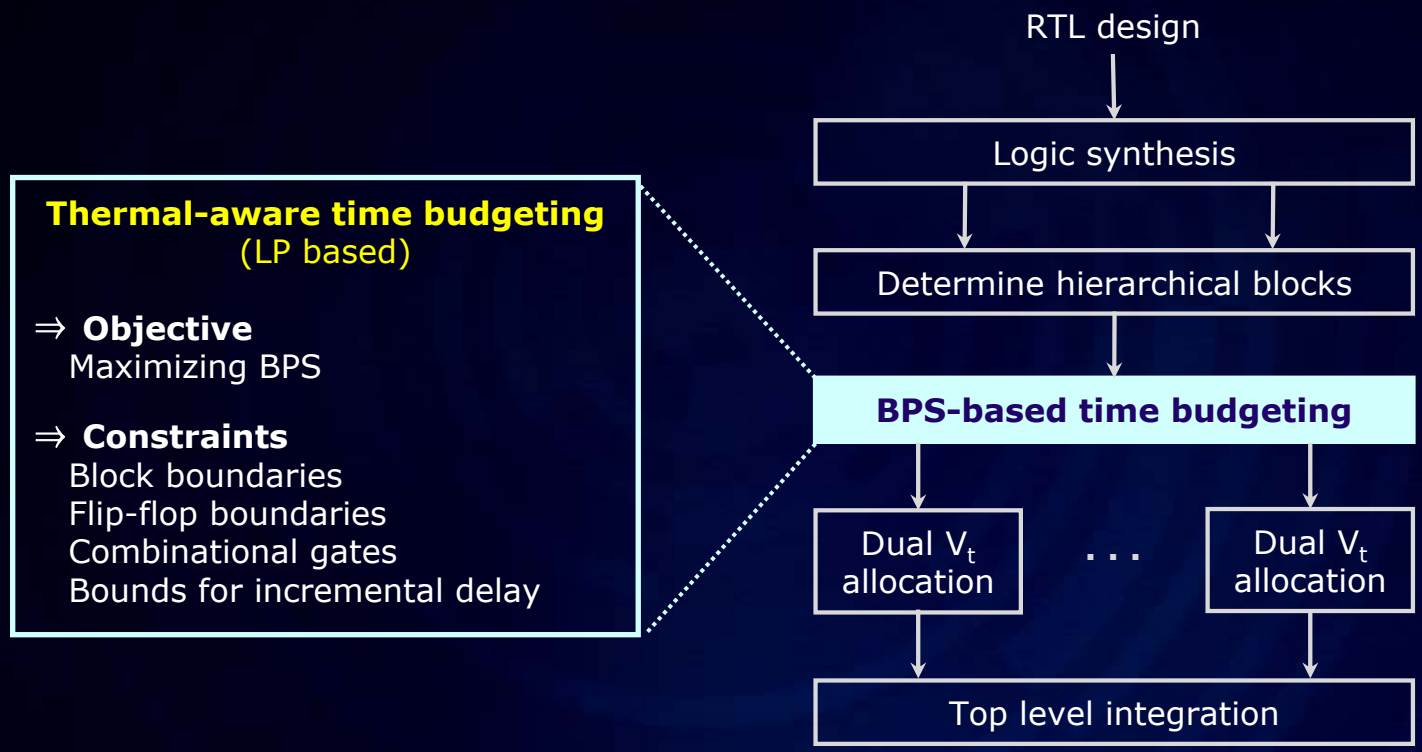
    - Bounds for incremental delay

$$0 \leq \Delta d_i \leq B_i$$

where $B_i = d_i\big|_{V_t=high} - d_i\big|_{V_t=low}$

BPS-Based Time Budgeting

- Overall design flow



Experimental Results

- Leakage power

Benchmark					Leakage power (uW)		BPS/Conv (X)
Name	# Gates	# FFs	# Blocks	P (ns)	Conventional	BPS-based	
b17	23063	1414	8	2.8	296.0	183.0	0.62
ac97	8756	2181	8	0.9	33.2	16.6	0.50
oc54	13481	1426	5	2.3	84.6	71.8	0.85
ps2	2255	254	4	0.9	13.5	9.8	0.73
ucore	14149	1202	5	1.7	35.1	25.3	0.72
Warp	21520	1670	7	2.0	124.9	108.0	0.87
s13207	2260	490	4	1.3	24.6	19.3	0.79
s38417	6899	1564	8	1.0	69.5	47.2	0.68
Average							0.72

Experimental Results

- Run time

Benchmark					Run time (s)				BPS/Conv (X)
					Conventional		BPS-based		
Name	# Gates	# FFs	# Blocks	P (ns)	Budgeting	Dual V_t	Budgeting	Dual V_t	
b17	23063	1414	8	2.8	98	2197	3477	1159	2.02
ac97	8756	2181	8	0.9	10	470	303	153	0.95
oc54	13481	1426	5	2.3	24	787	934	662	1.97
ps2	2255	254	4	0.9	4	43	12	32	0.94
ucore	14149	1202	5	1.7	14	395	892	221	2.72
Warp	21520	1670	7	2.0	38	1592	2311	1172	2.14
s13207	2260	490	4	1.3	4	42	11	26	0.8
s38417	6899	1564	8	1.0	12	425	250	270	1.19
Average									1.59

Conclusion

- BPS-based time budgeting
  - Bounded potential slack
    - Good metric for dual v_t allocation
  - LP based formulation
  - 30% leakage saving on average compared to conventional approach
  - However, 60% runtime increase on average compared to conventional approach