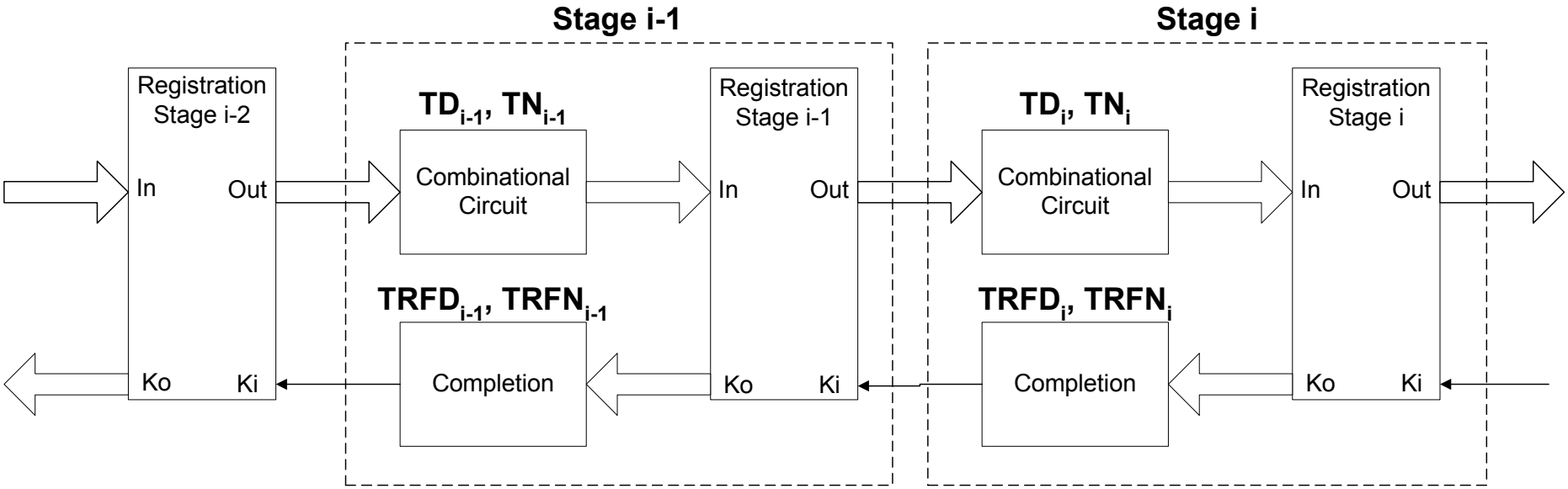
NCL Throughput Derivation



Stage	Sub-cycle	Initial State	Wavefronts								
			1	**2**	**3**	**4**	**5**	6	7	8	9
i-1	**D_{i-1}**		X				**X**				X
	N_{i-1}	X			**X**				X		
	RFD_{i-1}	X				**X**				X	
	RFN_{i-1}			**X**				X			
i	**D_i**			**X**				X			
	N_i	X				**X**				X	
	RFD_i	X					**X**				X
	RFN_i				**X**				X		

NCL Throughput Derivation (continued)

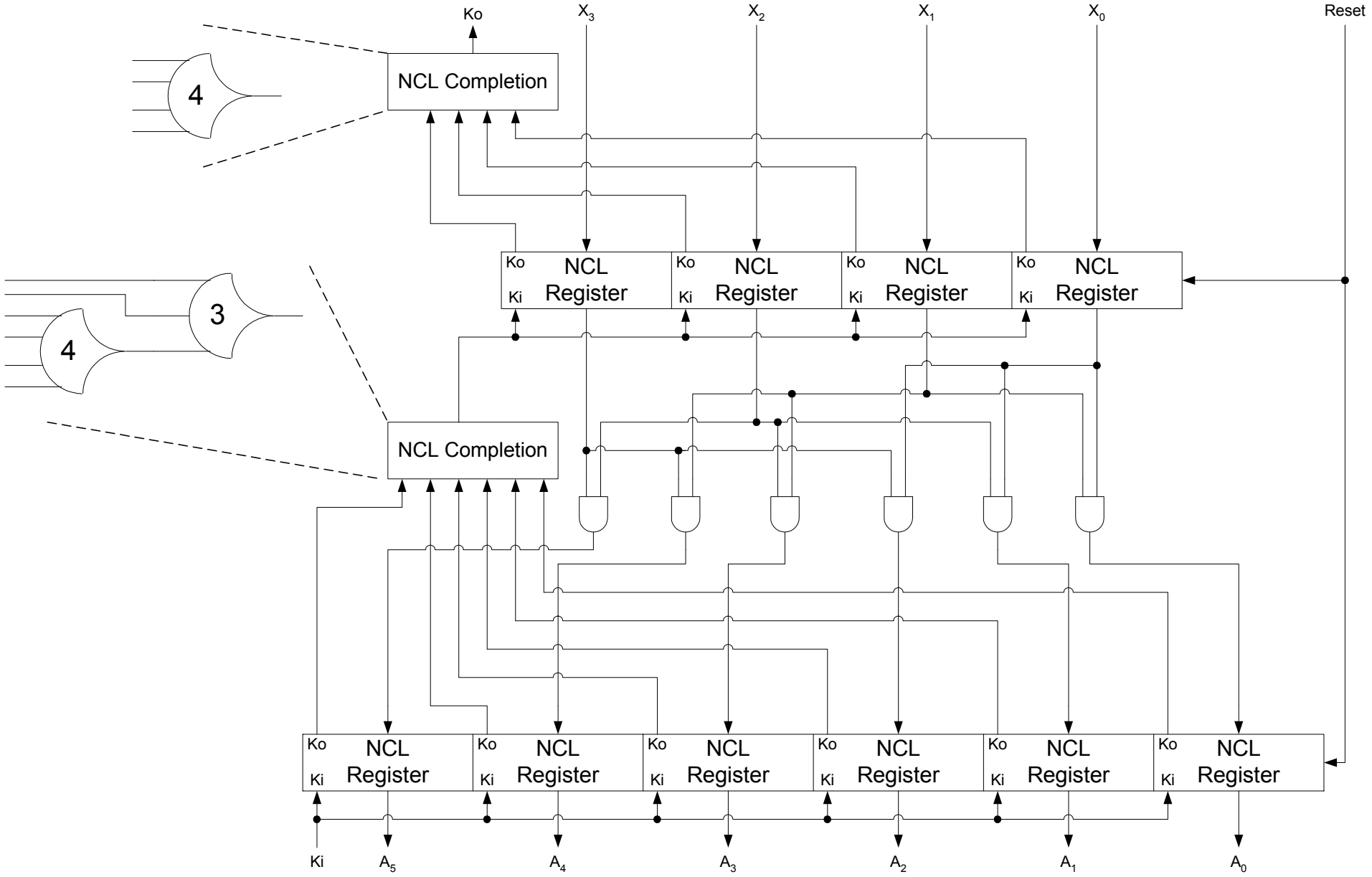
- $T_{DDi} = MAX(TD_i, TRFN_{i-1}) + MAX(TN_{i-1}, TRFN_i) + MAX(TN_i, TRFD_{i-1}) + MAX(TD_{i-1}, TRFD_i)$

$T_{DDmax} = TRFD_1 + TD_1 + TRFN_1 + TN_1$
for (i = 2 to N) loop
$T_{DDtemp} = MAX(TRFD_i + TD_i + TRFN_i + TN_i,$
$TRFD_{i-1} + TD_{i-1} + TD_i + TRFN_i,$
$TRFN_{i-1} + TN_{i-1} + TN_i + TRFD_i)$
$T_{DDmax} = MAX(T_{DDtemp}, T_{DDmax})$
end loop
throughput = $1 / T_{DDmax}$

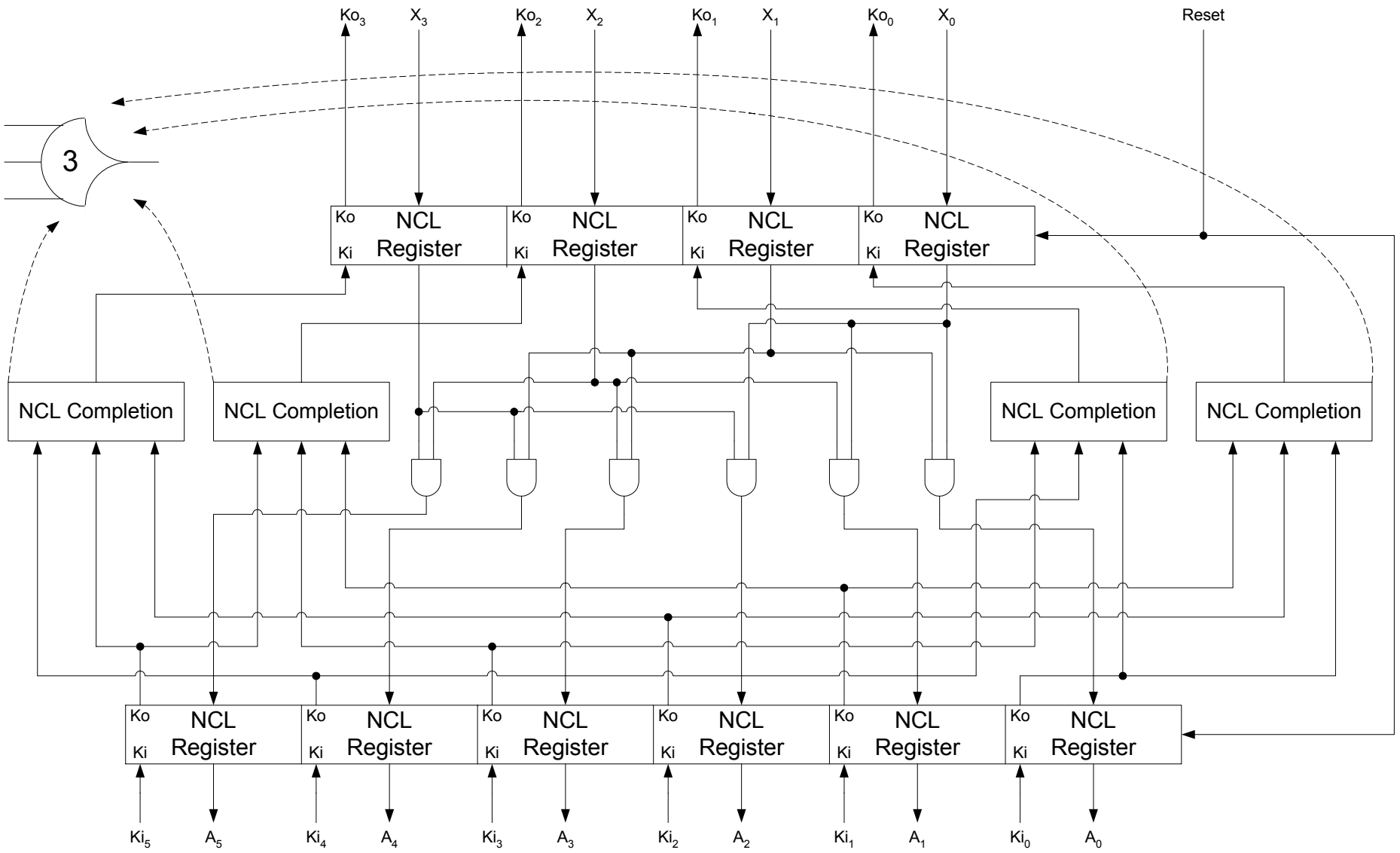
NCL T_{DD} Estimation

$T_{DDmax} = 2 \times (Dcomb_1 + Dcomp_1)$
for (i = 2 to N) loop
$T_{DDtemp} = 2 \times (Dcomb_i + Dcomp_i)$
$T_{DDmax} = MAX(T_{DDtemp}, T_{DDmax})$
end loop

Full-Word Completion

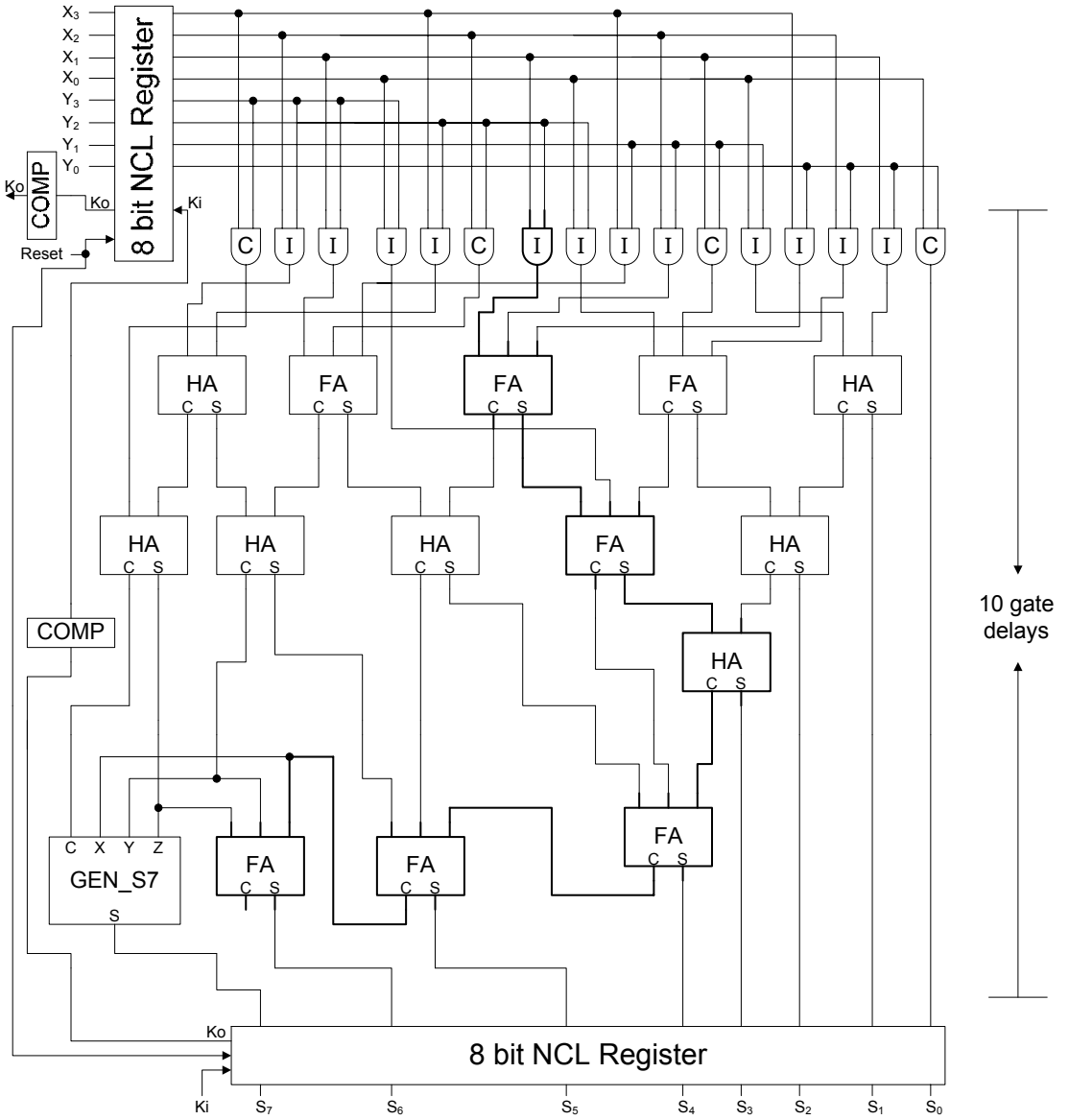


Bit-Wise Completion



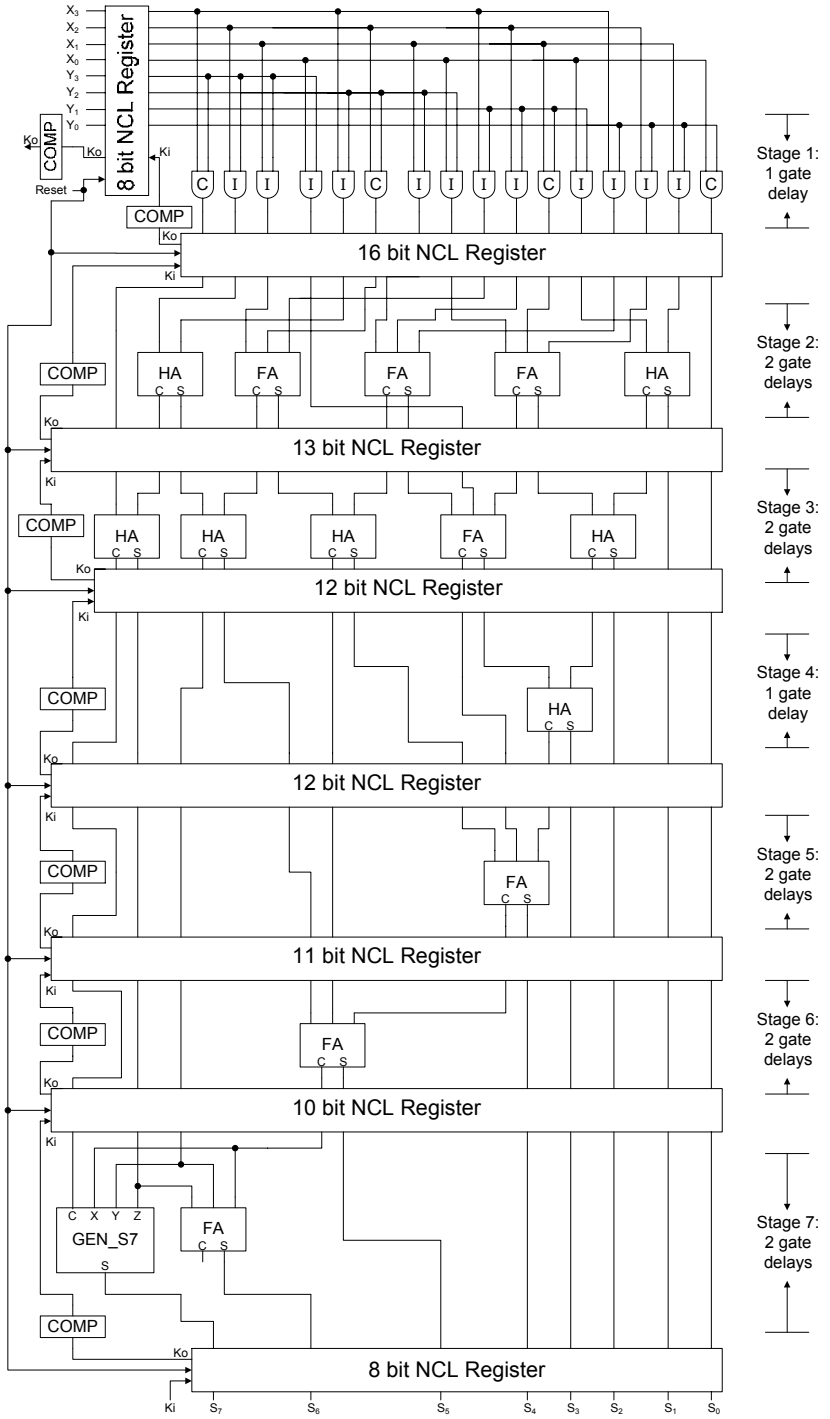
Non-Pipelined Multiplier

T_{DD} = 24 gate delays



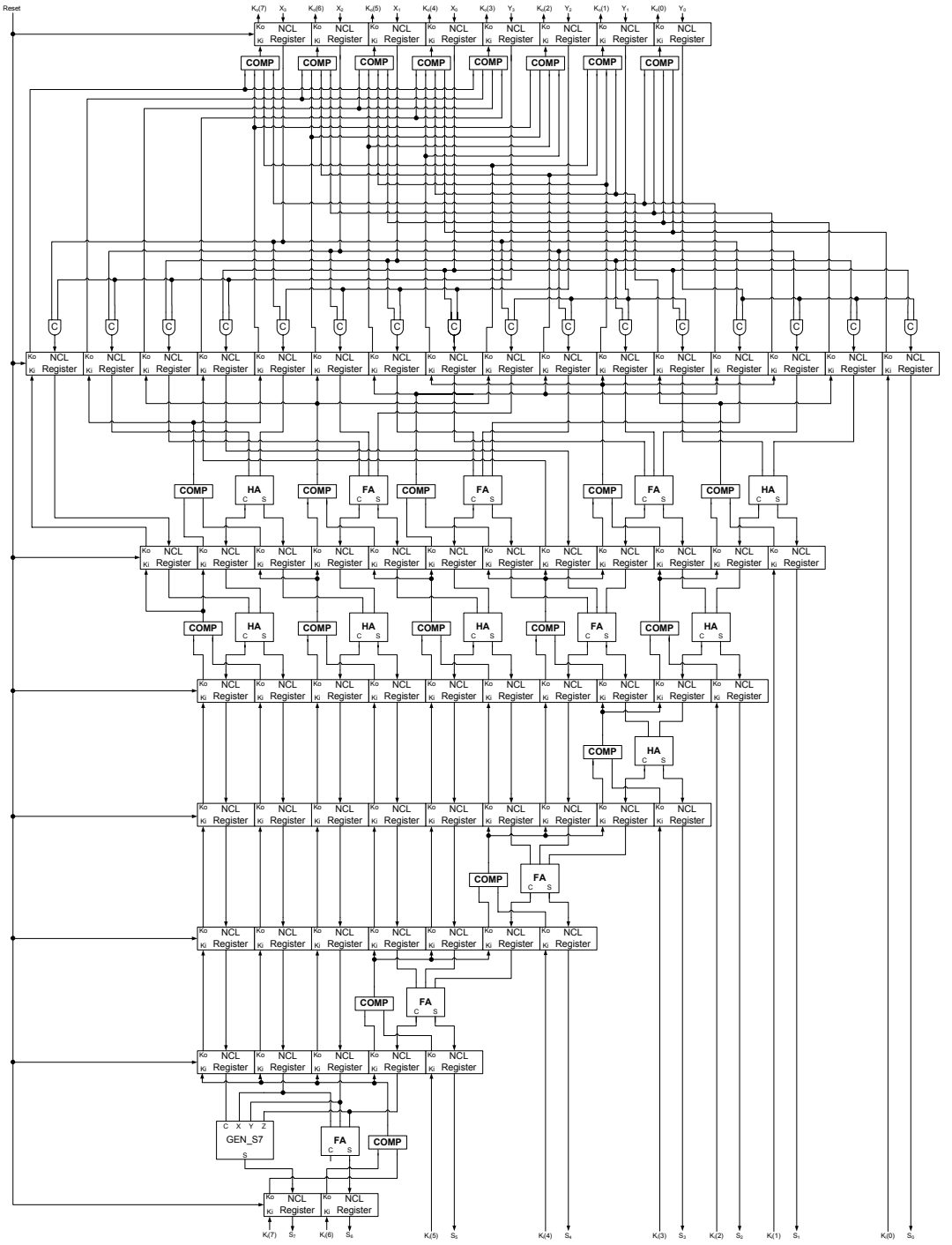
Full-Word Completion Multiplier

T_{DD} = 8 gate delays



Bit-Wise Completion Multiplier

T_{DD} = 6 gate delays



Multiplier Comparisons

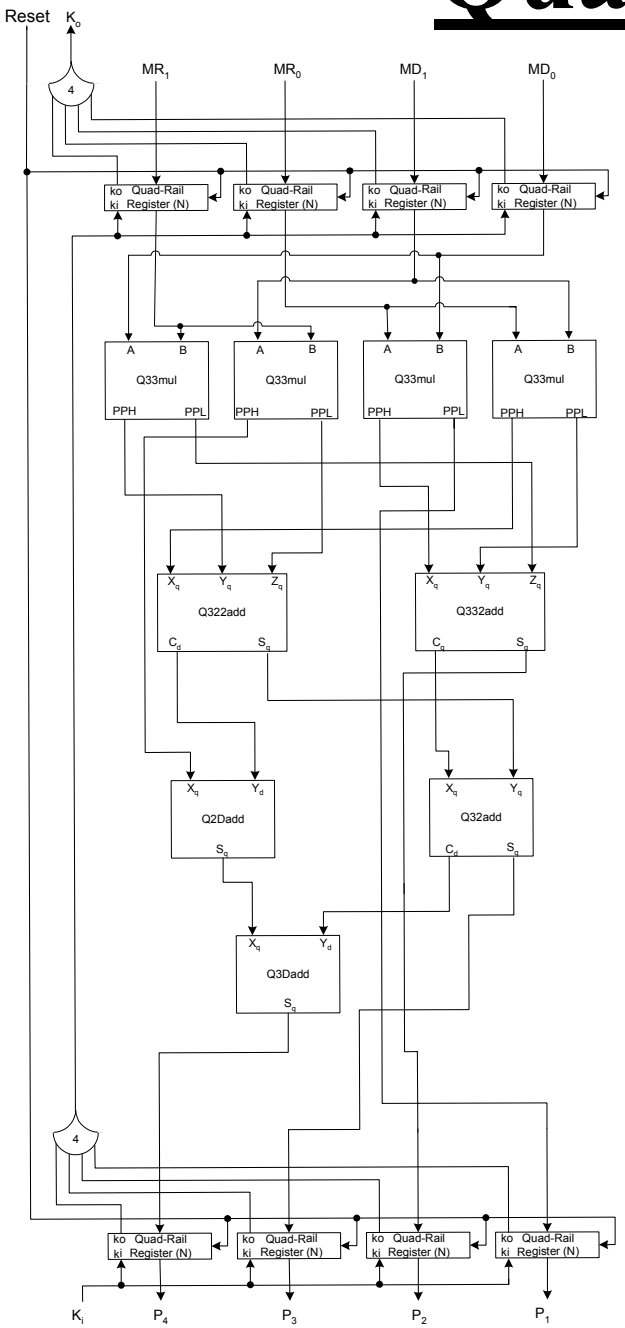
Multiplier Design	Maximum Combinational Delay per Stage (gate delays)	Maximum Completion Delay per Stage (gate delays)	Predicted Throughput $(\text{gate delays})^{-1}$	Simulated Throughput $(\text{ns})^{-1}$
Non-Pipelined	10	2	1/24 = 0.042	0.114
Full-Word	2	2	1/8 = 0.125	0.209
Bit-Wise	2	1	1/6 = 0.167	0.257

- Throughput Optimal Design Utilizes Bit-Wise Completion
  - speedup of 1.23 over Full-Word Completion Design
  - speedup of 2.25 over Non-Pipelined Design

NCL Pipelining Algorithm

1) initially partition an NCL circuit into stages of *primary components*
   - a primary component is defined as a component whose inputs only consist of the circuit's inputs or outputs of components that have already been added to a previous stage
2) calculate the combinational and completion delay for each stage and the maximum delay for the entire pipeline, utilizing both full-word and bit-wise completion strategies
3) merge stages to reduce latency and area, as long as doing so does not decrease throughput

Quad-Rail Multiplier



Component Type	Output Gate Delays	
	Carry / PPH	Sum / PPL
Q33mul	1	2
Q332add	3	3
Q322add	2	3
Q32add	2	2
Q2Dadd	N/A	1
Q3Dadd	N/A	1

Full-Word Completion Pipelining

Stage	D_{comb}	# Outputs	D_{comp}	delay
1	2	8	2	4
2	3	6	2	5
3	2	5	2	4
4	1	4	1	2
			max_delay	5

- Stages 3 and 4 can be combined: $D_{comb} = 3$, $D_{comp} = 1$, delay = 4

Bit-Wise Completion Pipelining

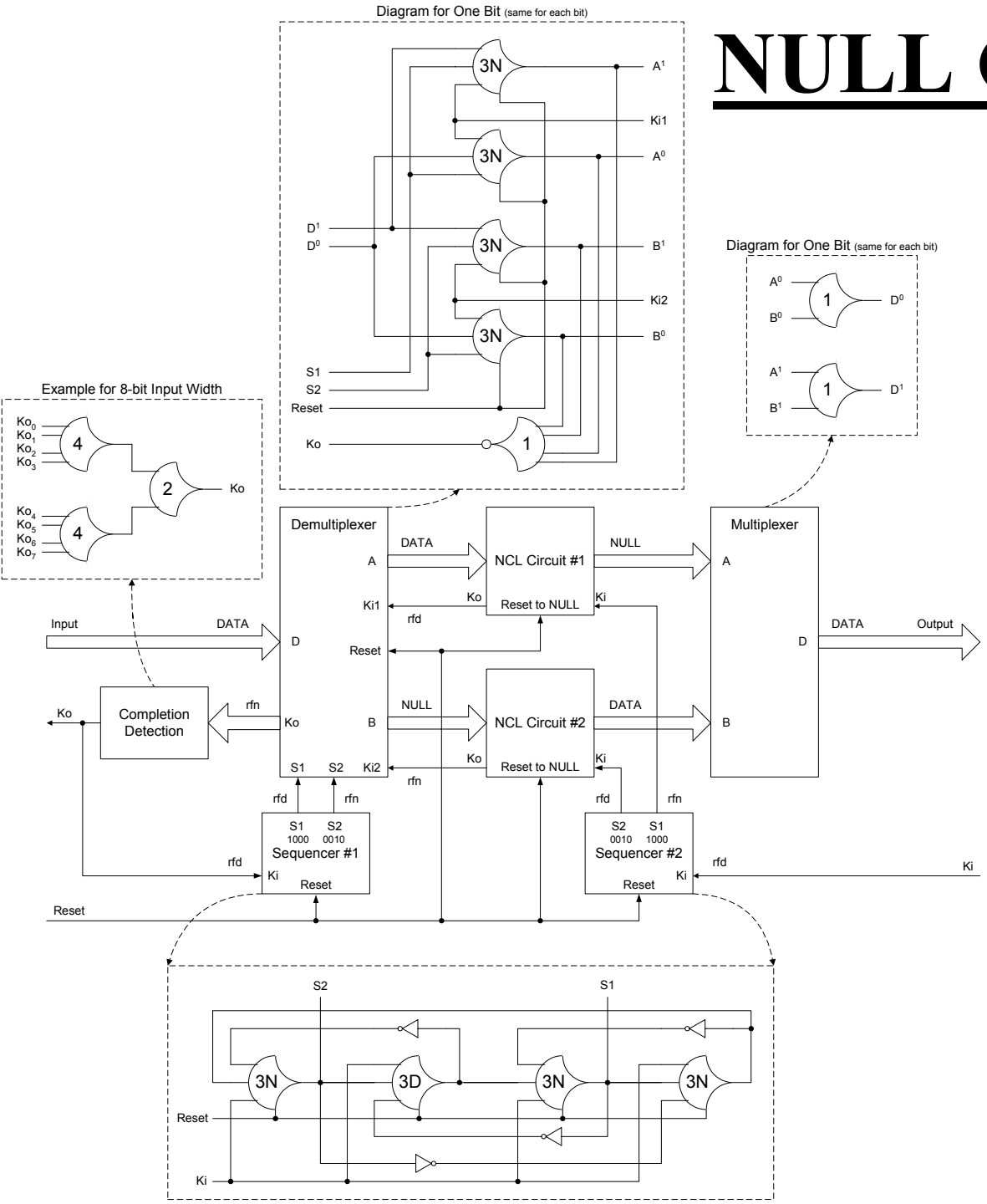
Stage	D_{comb}	max_outputs	D_{comp}	delay
1	2	4	1	3
2	3	2	1	4
3	2	2	1	3
4	1	1	0	1
			max_delay	4

- Stages 3 and 4 can be combined:

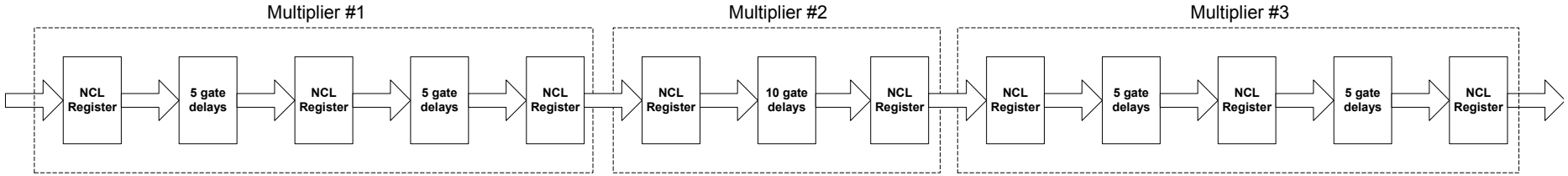
$D_{comb} = 3$, max_outputs $= 2 \rightarrow D_{comp} = 1$, delay = 4

NULL Cycle Reduction

- NCR applied to non-pipelined 4×4 multiplier
  - speedup of 1.61 achieved



NCR Applied to a Slow Pipeline Stage



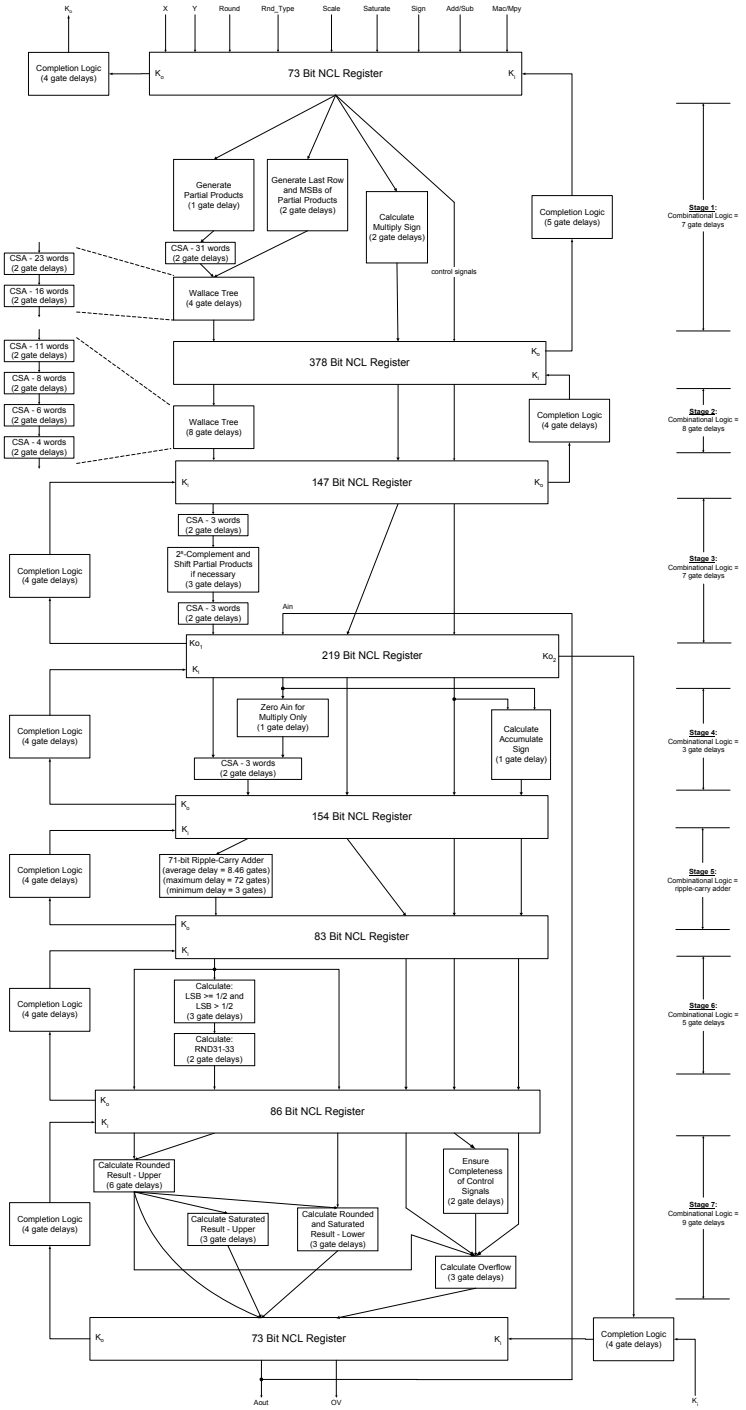
- Pipeline without NCR
  - T_{DD} = 8.42 ns
- Pipeline with NCR applied to Multiplier #2
  - T_{DD} = 6.96 ns
  - Speedup = 1.21

Pipelined MAC with Conditional Rounding, Scaling, and Saturation

Gate Count = 13,613

Minimum Feedback Path = 25 gate delays

T_{DD} = 12.7 ns



Pipelined MAC without Conditional Rounding, Scaling, and Saturation

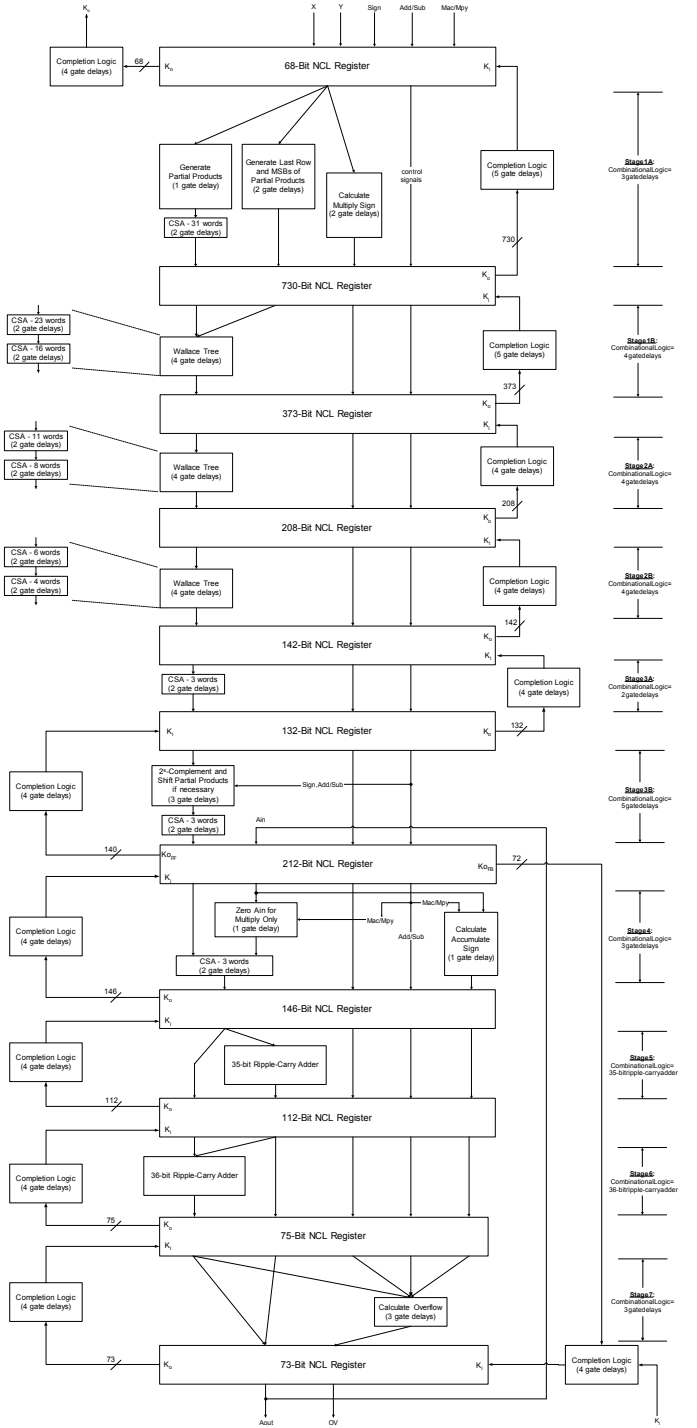
Gate Count = 16,169

Minimum Feedback Path = 18 gate delays

T_{DD} = 11.4 ns

Delay Analysis of 71-bit RCA

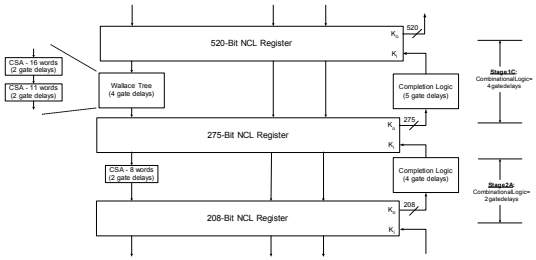
X	Percent ≥ X Gate Delays
11	21.8%
10	32.7%
9	41.4%
8	67.7%
7	91.4%
6	99.6%

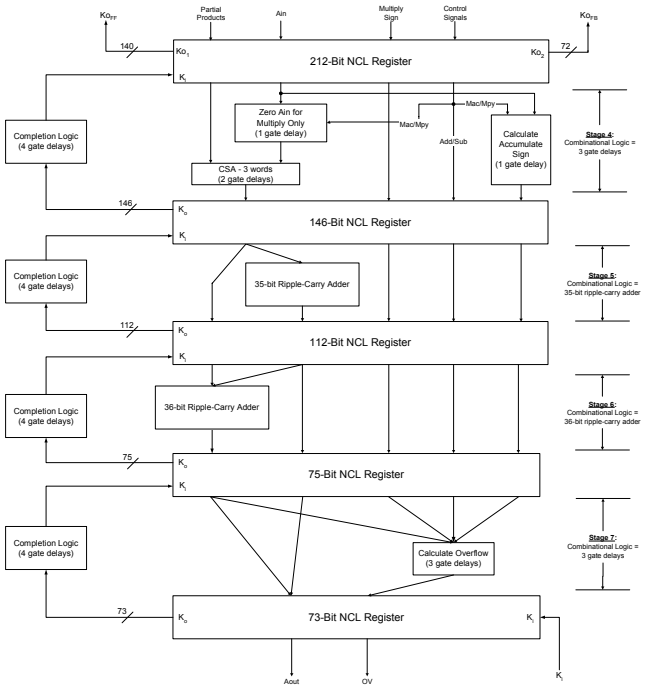


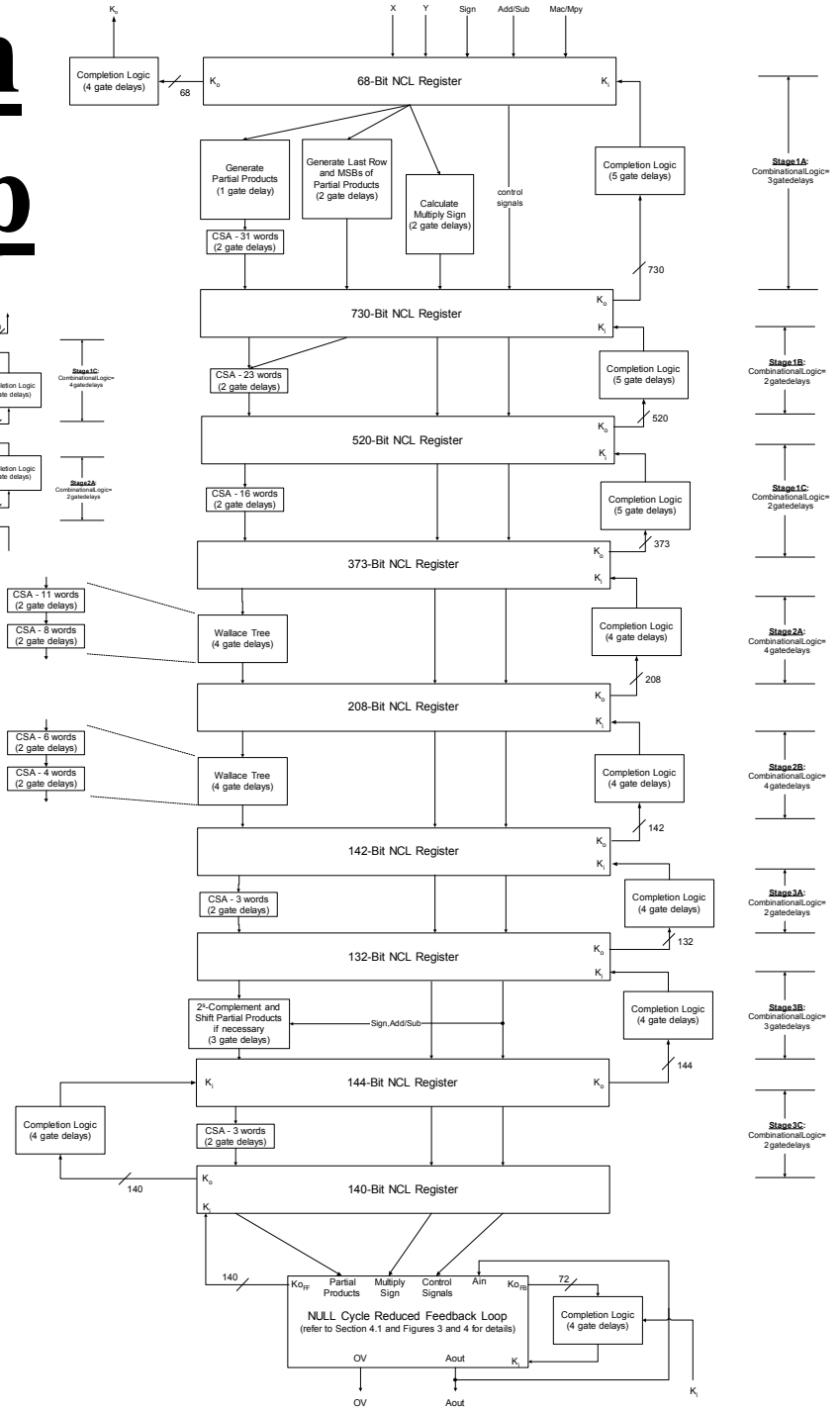
Pipelined MAC with NCR Feedback Loop

Gate Count = 22,804

T_{DD} = 8.7 ns



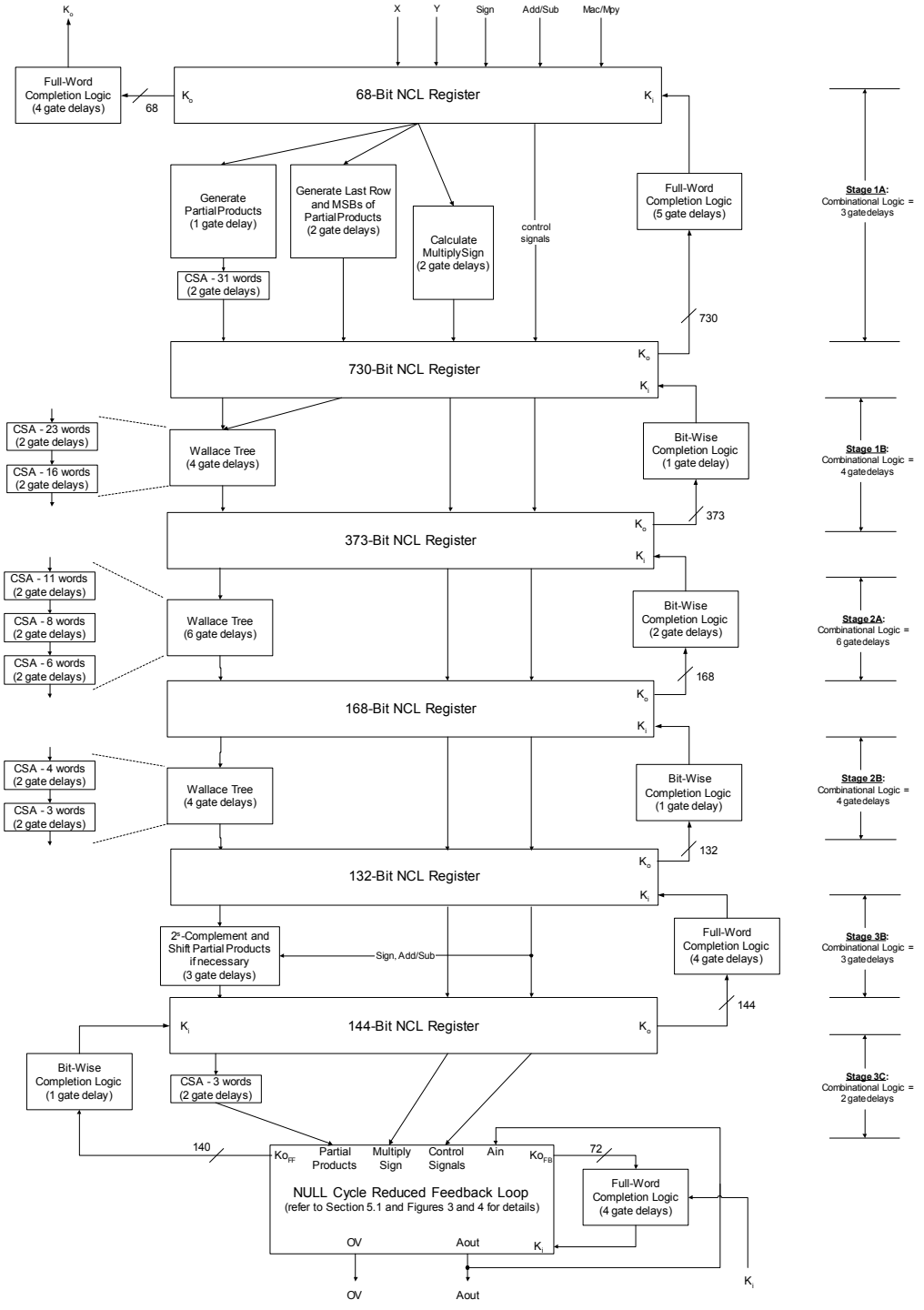




Pipelined MAC with NCR Feedback Loop and Bit-Wise Completion

Gate Count = 21,154

T_{DD} = 8.6 ns



MAC Comparisons

Design	Completion Strategy	T_{DD} (ns)	Gate Count
MAC with CRSS	Full-Word	12.7	13,613
MAC without CRSS	Full-Word	11.4	16,169
NCR MAC	Full-Word	8.7	22,804
NCR MAC	Bit-Wise	8.6	21,154

- Application of NCR to throughput-limiting feedback loop yielded speedup of 1.31