

ESE 3700 Project 2: A 16×4 SRAM in 22 nm PTM HP

Lucas Krippendorff

April 2026

1 Introduction

This report describes the design, validation, and characterization of a 16×4 single-port SRAM implemented in the 22 nm PTM high-performance BSIM4 process at $V_{DD} = 0.8$ V. The memory provides synchronous, full-word (4-bit) random read/write access using a 4-bit address, a single clock, and an active-high write enable. Correct operation is demonstrated across a suite of functional tests, and the design is characterized against the project figure of merit

$$\text{FOM} = 60 \cdot \text{BitcellArea} \cdot P \cdot D^2,$$

using the worst-case access delay and the average power during the specified all-zeros/all-ones write workload. The resulting FOM is $8.54 \times 10^{-30} \text{ m} \cdot \text{W} \cdot \text{s}^2$ at a 320 ps measurement period (3.125 GHz), well above the 500 MHz requirement.

2 Top-Level Architecture

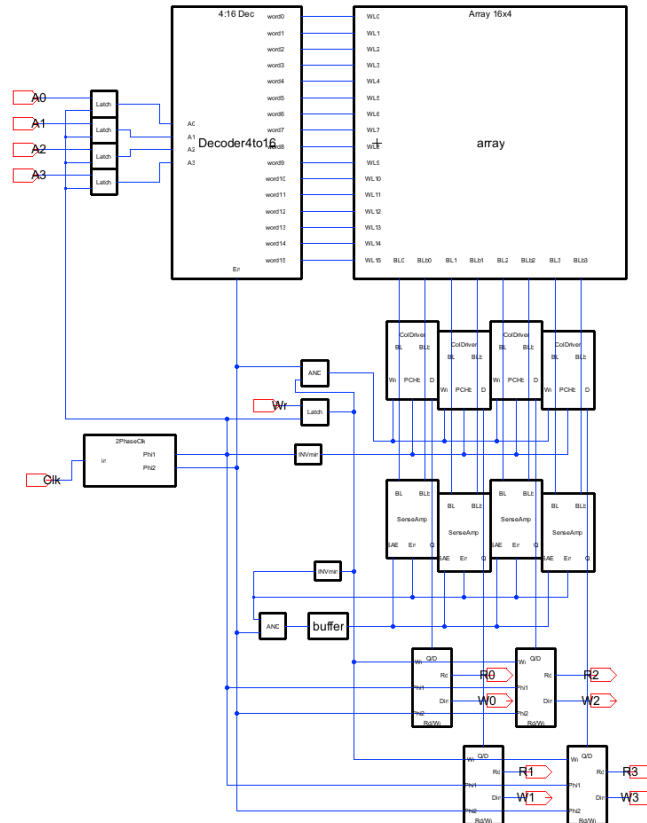


Figure 1: Top-level schematic (main). External pins: A0–A3, Clk, Wr, bit0–bit3. V_{DD}/GND are handled as globals through each leaf cell; they are not exported on the top icon.

The memory is hierarchically composed of the following sub-cells, each of which is described in Section 3:

- **2PhaseClock**: generates non-overlapping clocks ϕ_1 (precharge) and ϕ_2 (access) from the single input Clk.
- **Decoder4to16**: decodes A0–A3 and gates each row output with ϕ_2 to produce one of sixteen word-lines WL[0..15].
- **array**: 16 instances of wordRow, each containing four 6T Bitcells sharing a common WL.
- **ColumnDriver** $\times 4$: contains the precharge PMOS pair (gated by PCHb) and a tri-stated write driver (gated by Wr).
- **SenseAmp** $\times 4$: an isolated latch-type sense amplifier (Section 3.6).
- **ANDmin** + **buffer**: generates $SAE = \text{delayed}(\phi_2 \cdot \overline{Wr})$ and buffers it to all four SAs.
- **RdWr** $\times 4$: per-column bidirectional I/O block that merges D[i] and Q[i] onto a single shared bit[i] net (Section 3.7).

Control-signal routing. The external \overline{Wr} signal fans out to (1) the column-driver write enables, (2) an inverter producing \overline{Wr} which gates the sense-amp output tri-states, (3) one input of the SAE AND gate, and (4) the NMOS footer enable of each sense amp. The \overline{Wr} inverter is sized at $W = 4\times$ minimum to cover the combined load of roughly 620 nm of gate width on this network.

3 Sub-Component Design and Sizing

All sizings are expressed in nanometres of gate width W at fixed $L = 22$ nm. Every primitive uses a symmetric drawn width ($W_N = W_P$), so the unit size is 22 nm. A one-stage fan-out-of-four (FO4) delay in this process is approximately 10 ps at $V_{DD} = 0.8$ V, measured in a standalone inverter chain (/tmp/inv_fo1.spi).

3.1 Primitives

The periphery is built from a small library of sized CMOS primitives. Drawn widths are symmetric in the schematic; asymmetric drive is obtained by stepping up the instance size rather than by skewing W_P/W_N .

INV_min : $W_N=W_P=22$ nm ($1\times$ min). Used inside ANDmin, GateBasedLatch, the ϕ_2 -to-SAE delay chain, and as glue throughout Decoder4to16 and 2PhaseClock.

INV_sized : $W_N=W_P=88$ nm ($4\times$ min). Used as the PCHb driver, the \overline{Wr} driver, and the output stage of buffer.

INV_w8 : $W_N=W_P=176$ nm ($8\times$ min). Used as the final driver on ϕ_1 and ϕ_2 in 2PhaseClock, and as the last stage of buffer driving the four per-column SAE lines.

NAND2_min : $W_N=W_P=22$ nm. Used for per-row ϕ_2 -gating in Decoder4to16, as the AND inside ANDmin, and as the storage element of GateBasedLatch.

NOR2_min : $W_N=W_P=22$ nm. Used as the cross-coupled pair that enforces non-overlap in 2PhaseClock.

3.2 6T Bitcell

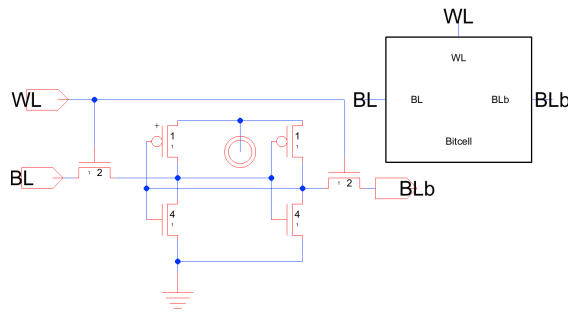


Figure 2: 6T SRAM bitcell. Width ratio PD:AX:PU = 4 : 2 : 1 (88 nm : 44 nm : 22 nm).

The storage cell is a six-transistor (6T) SRAM bitcell: a pair of cross-coupled CMOS inverters and two NMOS access transistors. Width ratios were chosen to satisfy the two competing static constraints:

Read stability. During a read the access transistor must not allow the precharged-high bitline to flip the internal low node. This requires the pulldown NMOS to be strong relative to the access NMOS. The *cell ratio* captures that relationship, $CR = PD/AX = 88/44 = 2$, so the low node rises only a few tens of millivolts during access, well below the inverter trip point.

Write-ability. During a write the access transistor must be able to pull the internal high node below the inverter trip point against the pullup PMOS. This requires the pullup PMOS to be *weak* relative to the access NMOS. The *pullup ratio* captures that relationship, $PR = PU/AX = 22/44 = 0.5$, so the access device wins cleanly.

Lecture 17 gives the thresholds for each ratio. The read-stability curve on slide 17 plots the voltage rise on the internal low node against CR; guaranteed read stability (the rise just reaching $V_{DD}/2$) sits near $CR \approx 1.2$, and the $CR = 2$ here is comfortably above that boundary. The write-ability curve on slide 24 plots the post-write cell voltage against PR; guaranteed writability sits near $PR \approx 1.8$, above which the pullup wins and the cell refuses to flip. The $PR = 0.5$ here is far below that boundary, so the access device wins cleanly and the new value reaches ground with margin.

Combined, the absolute ratio $PD:AX:PU = 4:2:1$ biases the cell toward read stability, a sensible trade-off given that reads happen on every cycle regardless of Wr . Summing the six device widths gives a total bitcell area of $2(88) + 2(44) + 2(22) = 308 \text{ nm}$ ($0.308 \mu\text{m}$), which is the quantity used in the FOM (Section 7.3).

3.3 Two-Phase Non-Overlapping Clock Generator

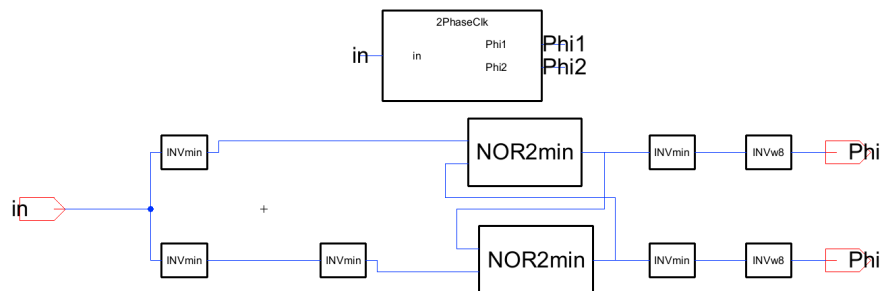


Figure 3: Two-phase non-overlapping clock generator. The cross-coupled NOR2_min guarantees that ϕ_1 and ϕ_2 are never simultaneously high.

A single external Clk is converted into two non-overlapping clocks: ϕ_1 is the precharge phase and ϕ_2 is the access phase. The non-overlap gap between ϕ_1 falling and ϕ_2 rising is the entire reason for using two phases, and is what guarantees that the precharge PMOS pair is fully off before any word-line rises.

Each output path uses the chain NOR2_min \rightarrow INV_min \rightarrow INV_sized so that the two inverters after the cross-coupled NOR preserve the intended polarity (a single inverter after the NOR would swap the ϕ_1/ϕ_2 labels). The final-stage driver on ϕ_2 is sized at $W = 8 \times$ minimum (INV_w8, $W_N=W_P=176$ nm) because it drives one NAND input per row, or 16 NAND inputs total (≈ 700 nm of gate width). The corresponding final-stage driver on ϕ_1 can remain minimum because it drives only one downstream inverter (the one that produces PCHb, discussed in Section 3.5).

3.4 4-to-16 Row Decoder

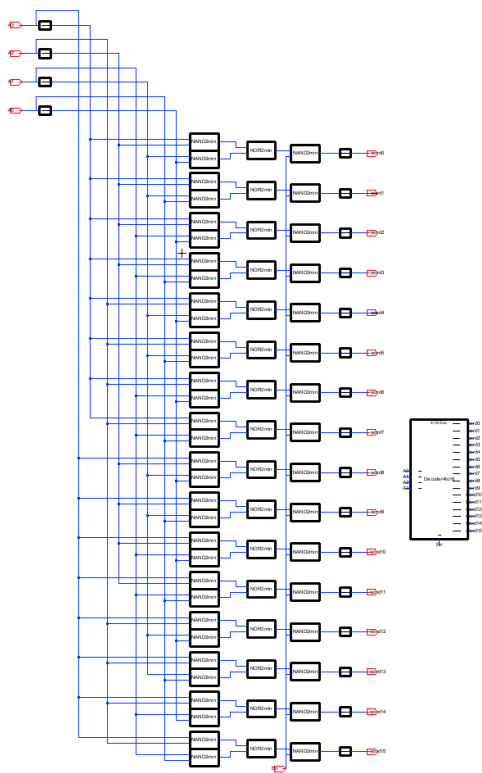


Figure 4: 4-to-16 row decoder. Each row output is $WL[i] = \overline{\text{dec}[i]} \cdot \phi_2$ followed by an inverter, giving the required positive-logic, ϕ_2 -gated word-line pulse.

The four address bits are decoded through an AND tree into 16 one-hot row lines $\text{dec}[i]$. Each row output is then gated with ϕ_2 through a NAND2_min and an INV_min, so $WL[i]$ is high only when (a) the address selects row i , and (b) the SRAM is in its access phase. The ϕ_2 gating is what makes changes to the address during the precharge phase harmless; WL cannot rise until the next access phase begins, and the precharge/access invariant (Section 4) keeps the cell contents undisturbed in the interim.

Each WL drives four access-transistor gates (one per column), for roughly 175 nm of gate capacitance, easily within the drive of the minimum inverter that terminates the decoder row.

3.5 Column Driver and Precharge

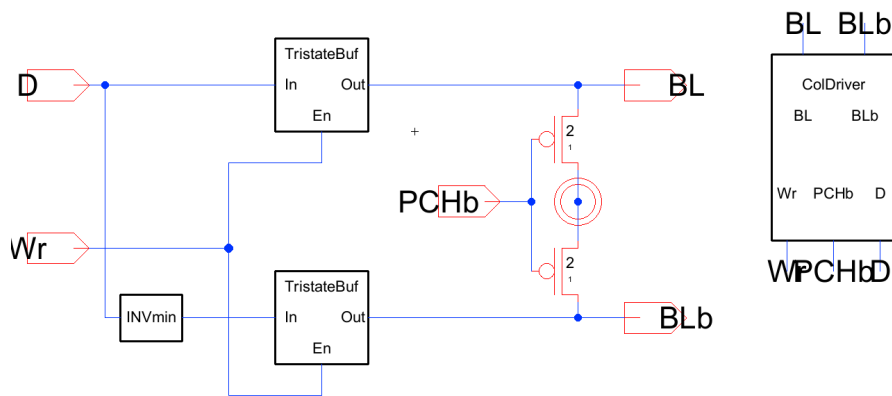


Figure 5: One of four identical column drivers. The two PMOS at the top are the precharge pair (gated by PCHb); the lower tri-state write path (RdWr) drives BL/BLb from the shared bus during write and goes high- Z during read.

Each column has two responsibilities that the column driver multiplexes in time:

1. **Precharge.** Two PMOS, one on BL and one on BLb, are gated by PCHb. When PCHb is low, both bitlines are pulled to V_{DD} ; when it rises, both PMOS turn off and the bitlines are held high by their own capacitance for the duration of the access. Each precharge PMOS is sized at $W \approx 176$ nm; the driving PCHb inverter (INV_sized) is sized at $W = 4\times$ to match the combined $1.4\ \mu\text{m}$ gate capacitance across four columns at roughly FO4.
2. **Write.** The write path takes bit[i] from the shared bus, drives it onto BL and its complement onto BLb, and tri-states when Wr is low. The tri-state is what allows the same BL net to be driven by the column driver during write and passively observed by the sense amp during read.

PCHb polarity: $\overline{\phi_1}$, not ϕ_2 . PCHb must be low during ϕ_1 and high during ϕ_2 , a polarity satisfied by either ϕ_2 or $\overline{\phi_1}$. I use $\overline{\phi_1}$ because it rises with the falling edge of ϕ_1 , one non-overlap gap before ϕ_2 rises, so precharge turns off before the write driver is enabled. Tying PCHb to ϕ_2 would erase that gap and let the precharge PMOS briefly fight the write driver on every write. On a write-1, for example, the driver must pull BLb down against a still-on precharge PMOS pulling it up.

3.6 Sense Amplifier

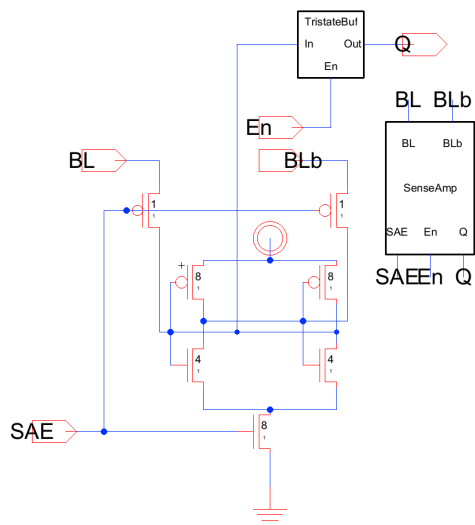


Figure 6: Isolated latch-type sense amplifier. Cross-coupled CMOS inverter pair with two PMOS isolation pass gates between the bitlines and the internal latch nodes, and an NMOS footer gated by SAE. Topology adapted from [1].

The sense amplifier is an *isolated* latch-type rather than a plain bitline-coupled latch. Two PMOS pass transistors, gated by SAE (active-low from the sense-amp’s perspective), connect BL and BLb to the internal latch nodes during signal development. When SAE rises, the pass PMOS turn off, isolating the latch from the heavy bitline capacitance, and the NMOS footer (gated by SAE through the internal En) turns on to fire the latch. The regeneration then operates on the small internal node capacitance alone, which is considerably faster and lower-energy than a footer-only latch that must drag the bitlines through a full transition.

Control gating and output tri-state. The sense amp is armed only during read: its En pin is tied to \overline{Wr} , so the footer is off during write and the latch cannot be disturbed by the column driver. Its output, which drives bit[i] on the shared bus, is tri-stated by \overline{Wr} so that the column driver wins the bus during write.

SAE generation. SAE is produced as $SAE = \text{delayed}(\phi_2 \cdot \overline{Wr})$ by an ANDmin followed by a buffer, with the delay supplied by a short chain of INV_min stages on the AND output. The target delay from ϕ_2 rise to SAE rise is 75–100 ps. At less than roughly 50 ps the bitline differential is too small for the sense amp to resolve reliably; at greater than 100 ps the bitline has already developed well past the SA’s offset and additional delay is wasted. Measured values are reported in Section 5.

3.7 Shared I/O Bus (RdWr)

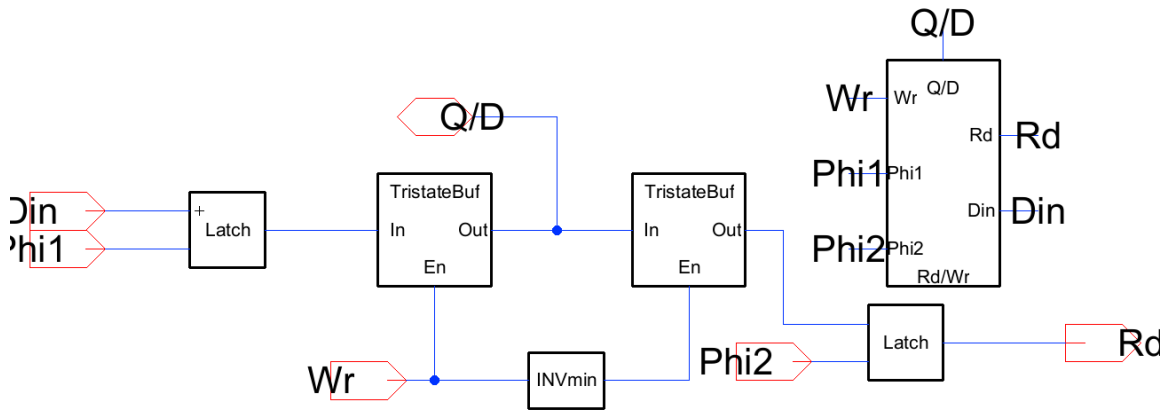


Figure 7: One of four per-column RdWr blocks. A ϕ_1 -clocked latch captures Din; its output drives the shared Q/D net through a TristateBuf enabled by Wr. A second TristateBuf, enabled by \overline{Wr} (via INVmin), drives a ϕ_2 -clocked latch that produces Rd.

To keep the external pinout small, each column's input $D[i]$ and sense-amp output $Q[i]$ share a single bidirectional net $bit[i]$. The write-side tri-state is enabled by Wr so the column driver wins the bus during write; the read-side tri-state is enabled by \overline{Wr} so the sense amp observes the bitlines undisturbed during read.

4 Timing Requirements for Correct Operation

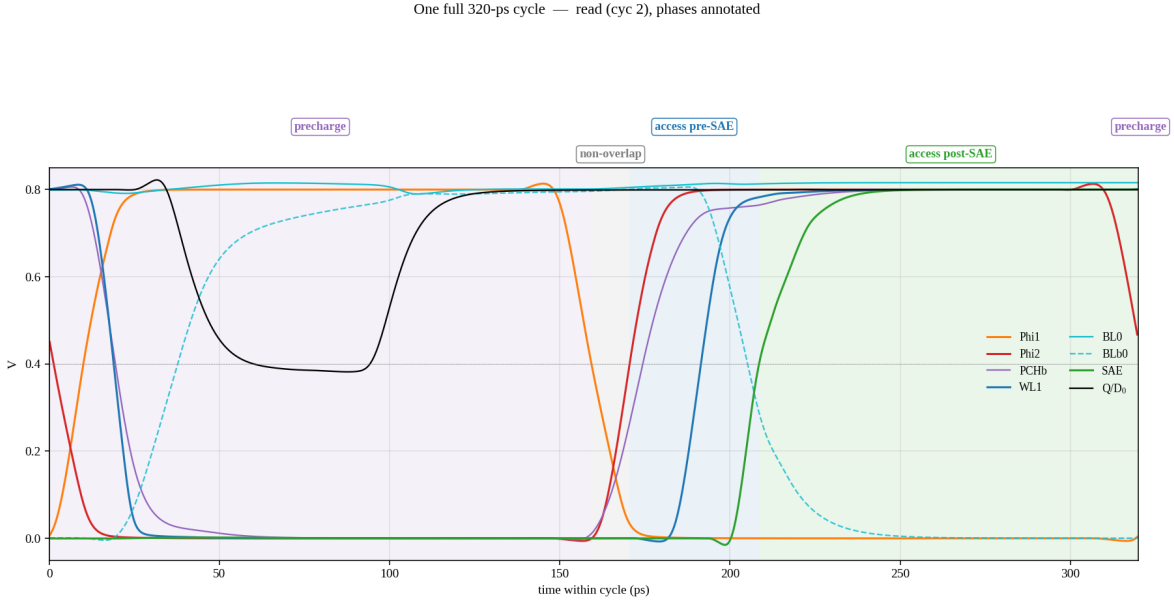


Figure 8: One 320 ps clock period showing the four phases (precharge, non-overlap gap, access pre-SAE, access post-SAE), overlaid with ϕ_1 , ϕ_2 , PCHb, WL, BL/BLb, SAE, and the internal sense-amp output Q .

Correct read and write operation is guaranteed by six invariants, each of which corresponds to a specific ordering of rising and falling edges inside one clock cycle:

1. Precharge completes before WL rises. Ensured because $\overline{\text{PCHb}} = \overline{\phi_1}$ rises with the falling edge of ϕ_1 , and WL depends on ϕ_2 , which rises only *after* the non-overlap gap.
2. WL rises before SAE fires. Ensured because SAE goes through the INV_min delay chain off ϕ_2 (target 75–100 ps), while WL goes through only a NAND2+INV (approximately 20 ps; see t_2 in Table 1).
3. SAE fires after the bitline has developed beyond the SA's offset. Ensured by tuning the SAE delay so the bitline differential at SAE rise exceeds 50 mV (Table 1, t_3 to t_4).
4. SAE falls before the next precharge begins. Automatic: SAE follows ϕ_2 and is further gated by $\overline{\text{Wr}}$.
5. Write driver is tri-stated during read. Enforced by the column driver's internal Wr gate.
6. Sense-amp output is tri-stated during write. Enforced by the SA's output $\overline{\text{Wr}}$ gate; also its footer is disabled so the latch cannot fight the column driver on BL/BLb.

These six conditions are all satisfied for any non-overlap gap greater than the worst-case precharge-off delay, any SAE delay greater than the bitline development time to 50 mV, and any clock period greater than T_{\min} . The measurement in Section 7.1 demonstrates that $T_{\min} = 298$ ps.

5 Read and Write Operations with Critical Path

The critical path for both operations is the access-phase chain from the rising edge of ϕ_2 to either the read output reaching $V_{DD}/2$ (read) or the cell's internal storage node reaching $V_{DD}/2$ (write). The waveforms below come from a functional test that writes 1111 to word-line 1 (address 0000), reads it back, writes 0000 to the same address, and reads it again. For the read endpoint, the implementation exports a dedicated per-column read output and that signal is measured directly. For the write endpoint, the measured node is the internal storage node of the cross-coupled inverter pair in the bitcell rather than the bitline, because the bitline can reach its final value before the latch has actually flipped.

5.1 Read Operation

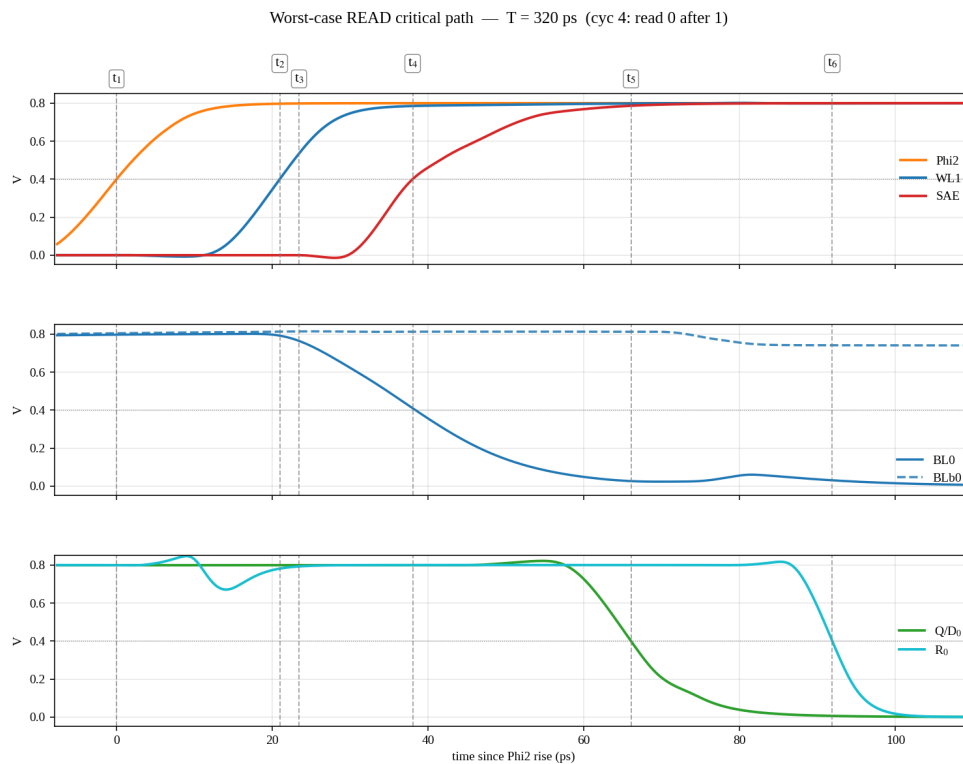


Figure 9: Worst-case read (cycle 4: read 0 after previously written 1). Markers $t_1 \dots t_6$ correspond to the stages in Table 1.

Table 1: Read critical-path breakdown at $T = 320$ ps. Offsets are from the rising edge of ϕ_2 . Threshold convention: $V_{DD}/2 = 0.4$ V for logic edges, 50 mV for bitline differential.

Mark	Offset (ps)	Event	Stage delay (ps)
t_1	0.0	ϕ_2 rises, access begins	
t_2	20.9	WL1 rises, decoder NAND+INV propagation done	20.9
t_3	23.4	BL differential reaches 50 mV	2.5
t_4	38.0	SAE fires, SA armed	14.6
t_5	66.0	Q/D_0 at $V_{DD}/2$, SA resolved	28.0
t_6	91.9	R_0 at $V_{DD}/2$, read latched	25.9

Total worst-case read delay: $D_{\text{read}} = 91.9$ ps.

5.2 Write Operation

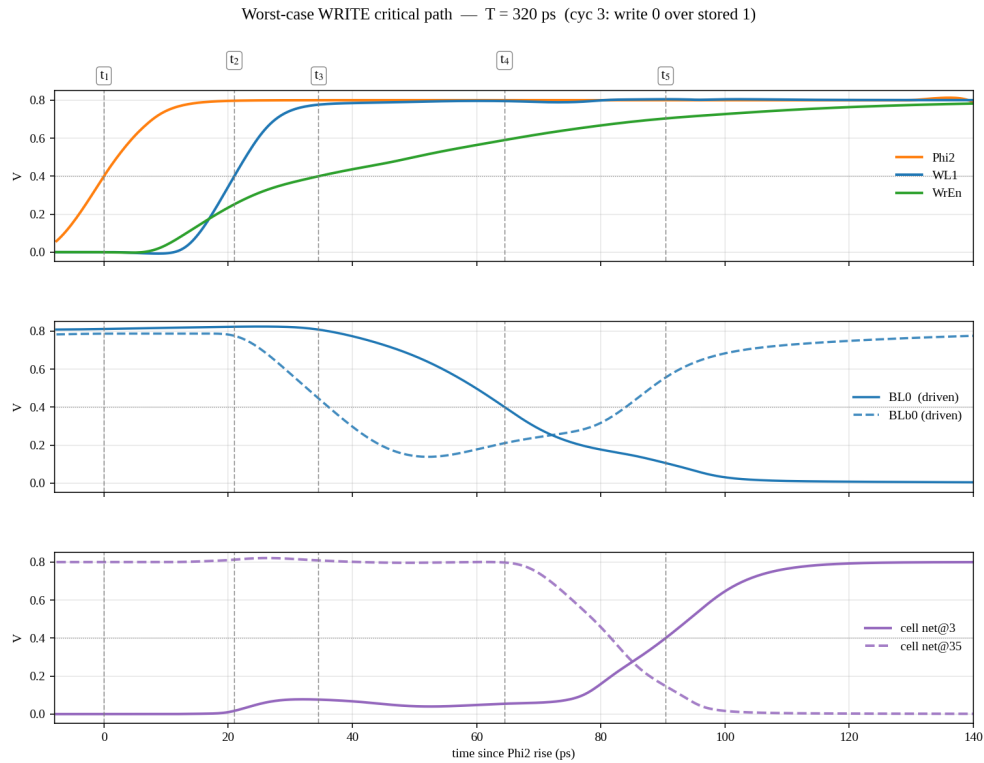


Figure 10: Worst-case write (cycle 3: write 0 over a previously stored 1). Markers $t_1 \dots t_5$ correspond to the stages in Table 2.

Table 2: Write critical-path breakdown at $T = 320$ ps.

Mark	Offset (ps)	Event	Stage delay (ps)
t_1	0.0	ϕ_2 rises, access begins	
t_2	21.0	WL1 rises	21.0
t_3	34.6	WrEn rises, column drivers enabled	13.6
t_4	64.5	BL0 crosses $V_{DD}/2$, data on bitline	29.9
t_5	90.5	Cell internal node at $V_{DD}/2$, write committed	26.0

Total worst-case write delay: $D_{\text{write}} = 90.5$ ps.

Worst-case delay (reported). $D = \max(D_{\text{read}}, D_{\text{write}}) = 91.9$ ps.

6 Design Validation and Test Cases

Correctness was validated bottom-up, starting from the bitcell and moving outward.

6.1 Bitcell Unit Test

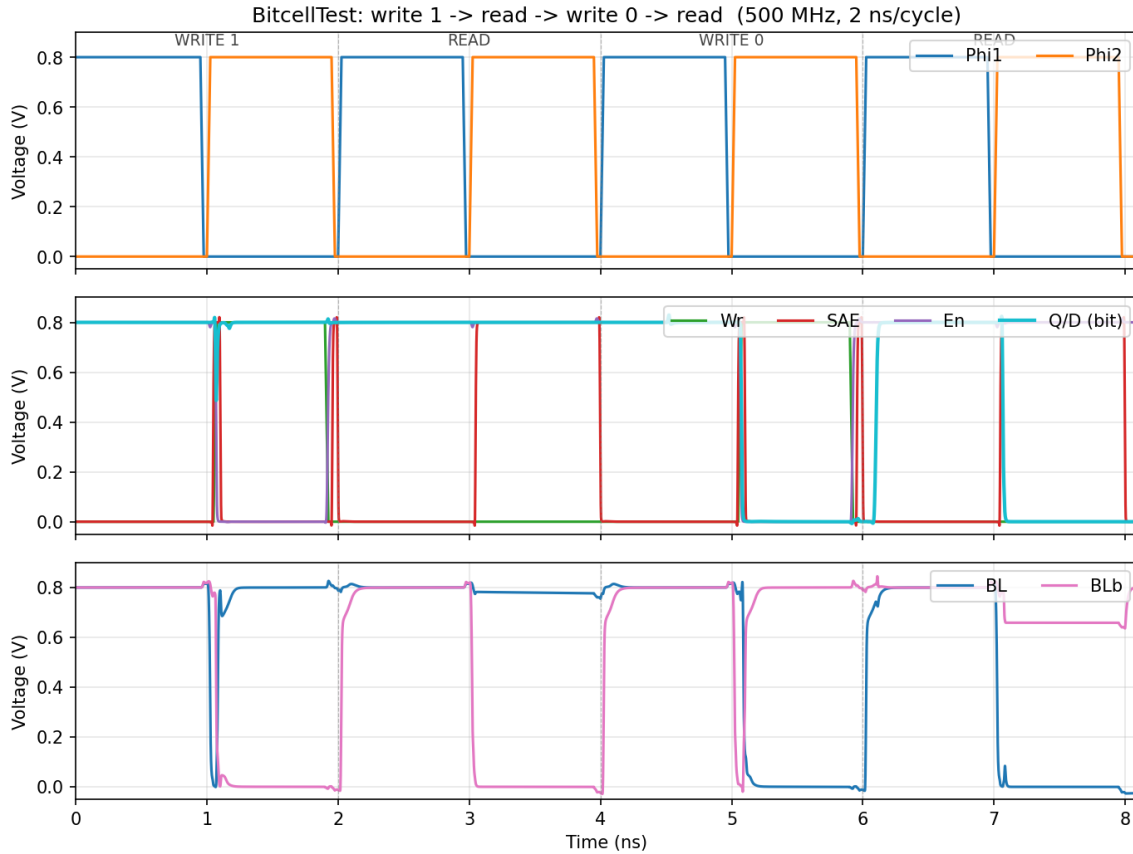


Figure 11: Bitcell unit test. A single cell is exercised by a PWL stimulus covering write 0, write 1, and a non-destructive read of each polarity. The internal storage nodes Q and \bar{Q} track the written value and remain stable during the subsequent read.

Figure 11 shows four back-to-back cycles at 2 ns per period, each exercising one operation. The waveform verifies correct behavior directly:

- Cycle 1, write 1 (0–2 ns). W_r is high and Q/D is driven to 1. During ϕ_2 (1–2 ns), BL holds high while BLb is pulled to 0, forcing the cell to store 1.
- Cycle 2, read 1 (2–4 ns). W_r is low. During ϕ_2 (3–4 ns), SAE rises and BLb drops (not BL), indicating the access transistors are discharging from the side holding 0; Q/D latches to 1, matching the stored value.
- Cycle 3, write 0 (4–6 ns). W_r is high and Q/D is driven to 0. BL is pulled to 0 while BLb holds high (the opposite of cycle 1), forcing the cell to store 0.
- Cycle 4, read 0 (6–8 ns). W_r is low. During ϕ_2 (7–8 ns), both bitlines fall slightly from the precharged-high state, but BL drops more than BLb (the opposite of cycle 2), so the sense amp latches Q/D to 0.

Together these four cycles verify both correctness conditions: the cell can be written in either direction (cycles 1 and 3), and a read in either polarity retrieves the stored value without disturbing it (cycles 2 and 4).

6.2 Full-Array Functional Test (mainTest1)

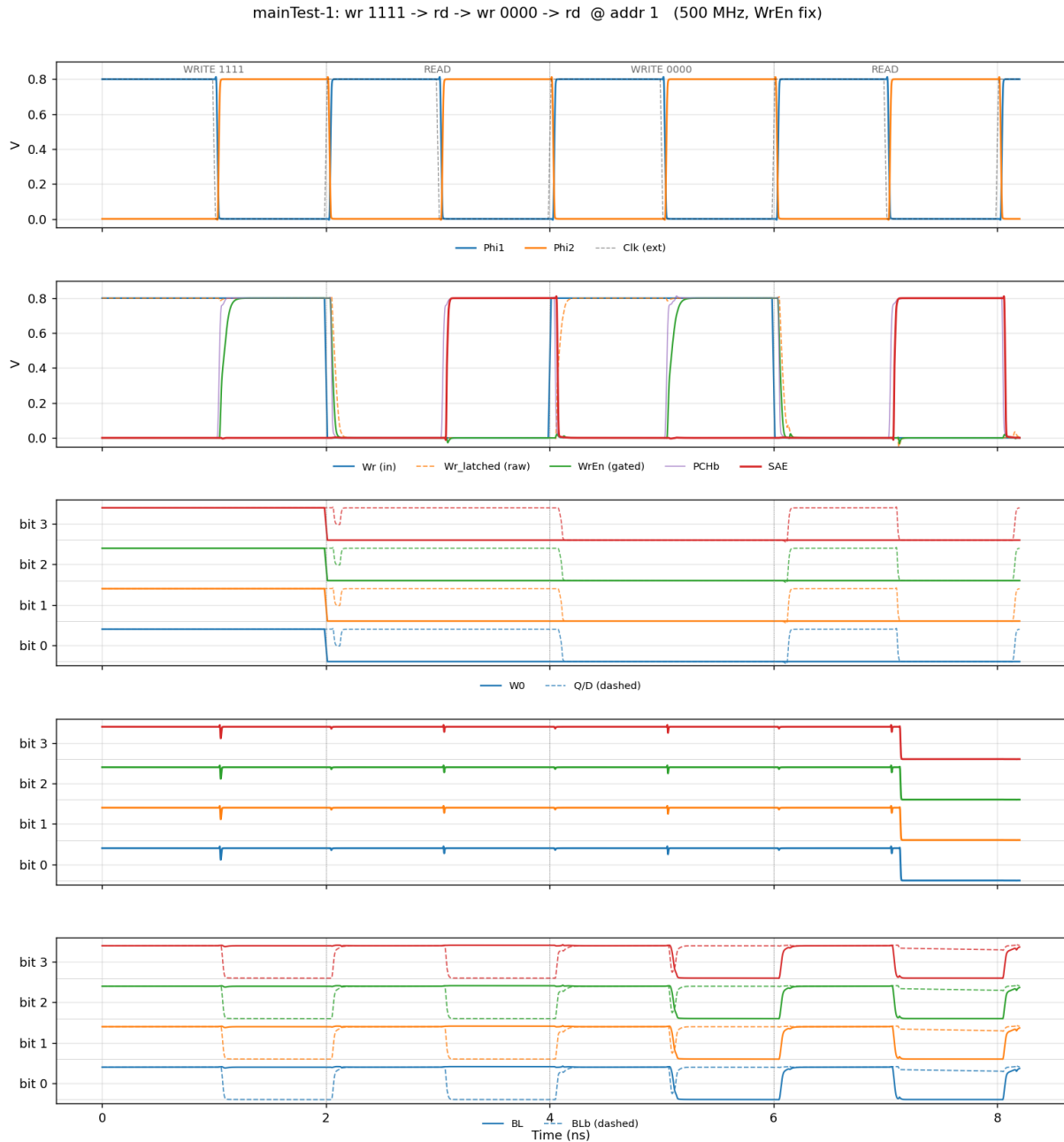


Figure 12: mainTest1: 1111 is written and read back, then 0000 is written and read back, all at the single address 0000. The plot does not label the rows, but the third plot from the top is the 4-bit write input Wr and the fourth is the 4-bit read output Rd.

Figure 12 is essentially the bitcell walkthrough from Section 6.1 copy-pasted four times in parallel, except now it is running on the full implementation, so the decoder, two-phase clock, and shared read/write bus are all in the loop. Every read matches the value that was just written.

6.3 Multi-Address Functional Test (mainTest2)

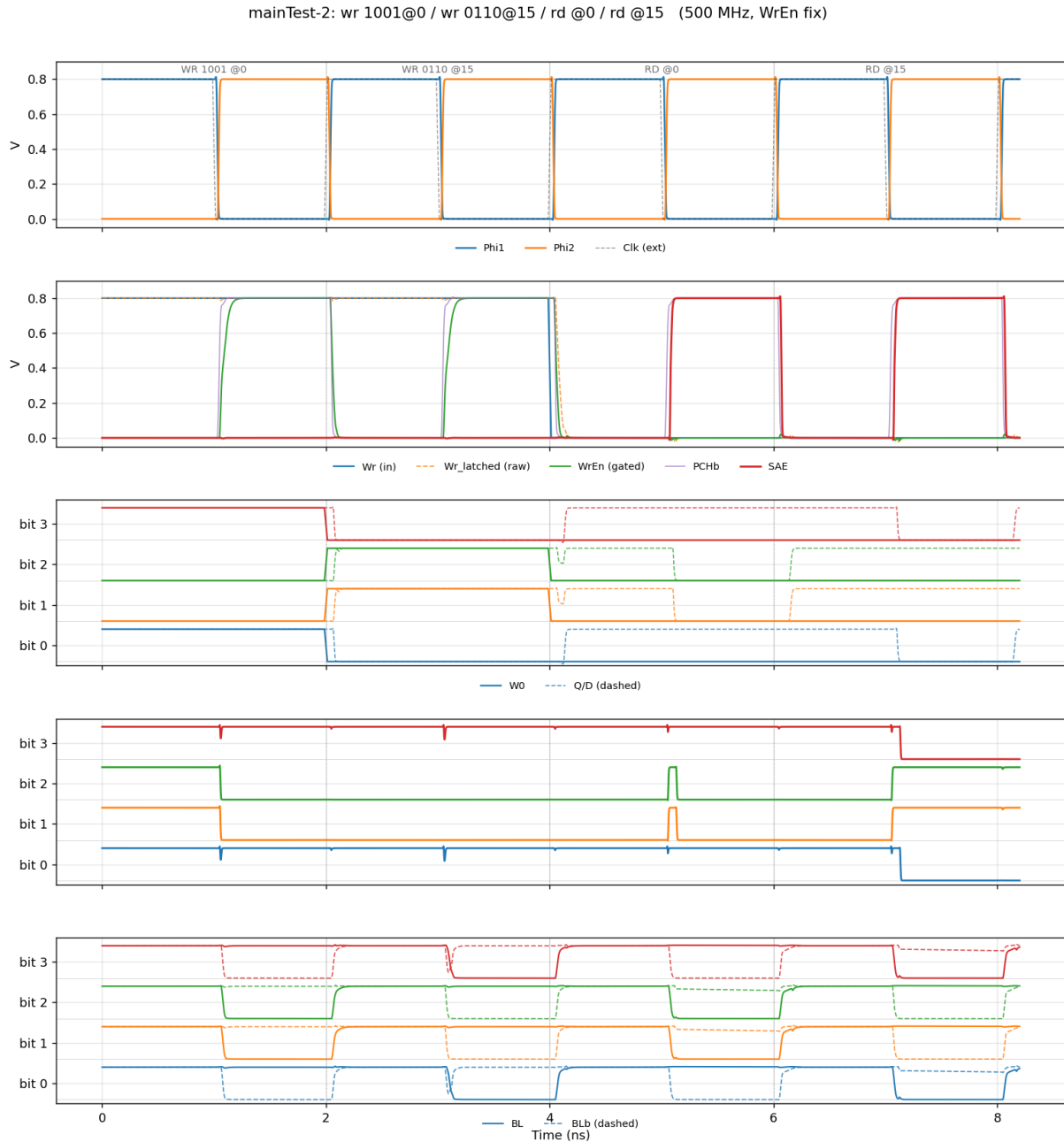


Figure 13: mainTest2: 1001 is written to address 0000 and 0110 is written to address 1111, then each is read back. Same row convention as Figure 12: third plot is Wr, fourth plot is Rd.

mainTest2 is the more interesting test because it checks that writes to different addresses do not disturb each other. At 0–2 ns the input 1001 is written to address 0000, and at 2–4 ns 0110 is written to address 1111. Reading address 0000 at 5–6 ns returns 1001 on the read output, and reading address 1111 at 7–8 ns returns 0110. Both words come back exactly as written, so storing one word does not corrupt the other.

6.4 Test-Case Coverage

The test suite exercises the following cases, which together cover every functional pathway of the design:

- Write/read both polarities at one address: mainTest1; detects any gross write-path or read-path malfunction.
- Multi-address data independence: mainTest2; two different 4-bit words written to two different addresses and read back separately, confirming the decoder routes the correct row and that stored data does not couple between rows.
- Worst-direction transitions: write 0 over stored 1 and read 0 after stored 1 are the slowest operations in each category (the SA starts with the precharged-high BL assisting a read 1 but fighting a read 0, and the column driver must pull BL from V_{DD} to ground for a write 0). These are the cycles plotted in Section 5.

7 Performance Characterization

7.1 Minimum Clock Period

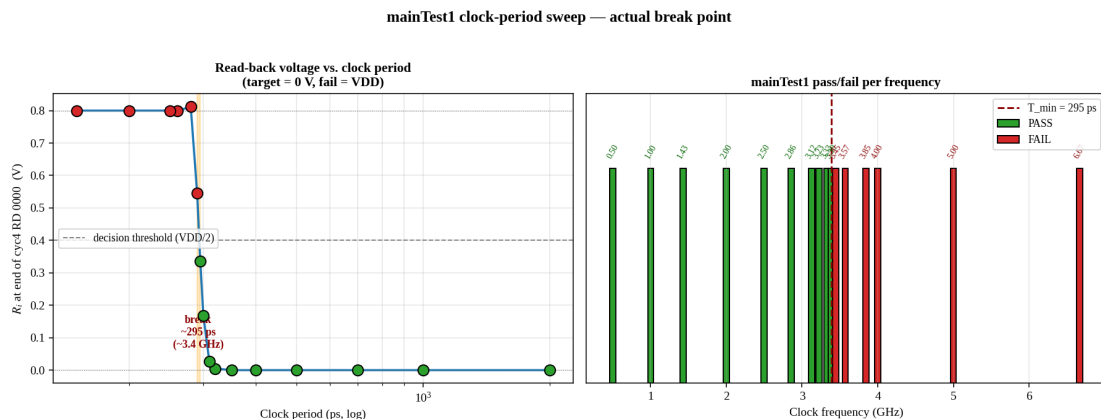


Figure 14: Binary search for the minimum clock period on the mainTest2 workload. $T_{\min} = 298$ ps (296 ps fails, 298 ps passes). The measurement in Section 7.2 uses $T = 320$ ps, roughly 7% margin above T_{\min} .

The minimum clock period was found with a binary search on the mainTest2 workload: the period was bracketed between a known-passing and a known-failing value and then halved on each iteration based on whether all reads matched. The search converged on $T_{\min} = 298$ ps (296 ps fails, 298 ps passes). The maximum operating frequency is therefore $f_{\max} = 1/T_{\min} = 1/(298 \text{ ps}) \approx 3.36 \text{ GHz}$, well above the project’s 500 MHz minimum.

At T_{\min} the read path still completes in approximately 91.7 ps, so the failure boundary is set by the *write* path (word-line pulse width and cross-coupled cell flip margin), not by sense-amp resolution. Adding margin on the write side would close this gap further.

7.2 Power Measurement

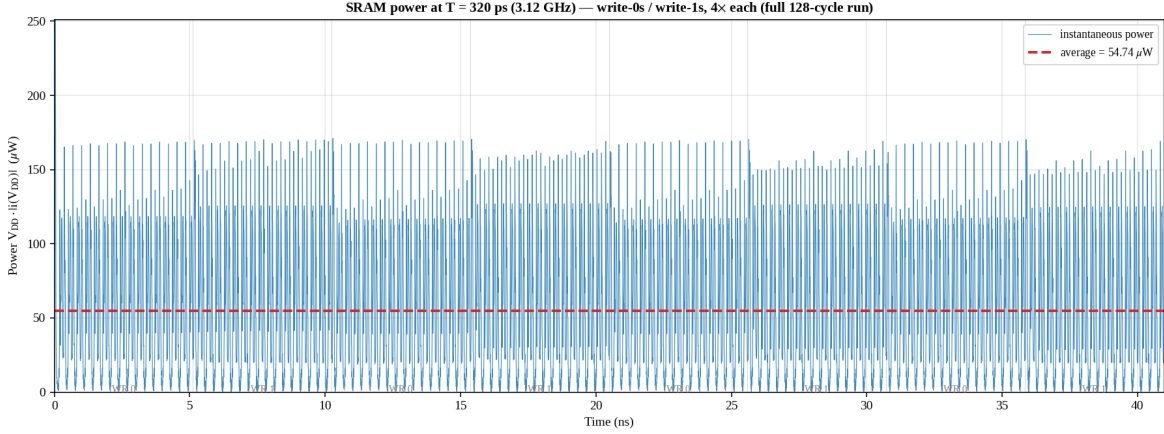


Figure 15: Instantaneous and time-averaged supply power at $T = 320$ ps during the FOM workload. The full simulation spans 128 clock cycles (16 addresses \times 2 patterns \times 4 repeats = 128 cycles \times 320 ps \approx 41 ns), consisting of four back-to-back repetitions of (write all-0s to every address, write all-1s to every address). The red dashed line is the arithmetic mean over the entire run: $P = 54.74 \mu\text{W}$.

Per the project specification, average power is measured while repeatedly writing all-0s and then all-1s to the entire array, with the clock set just above T_{\min} . I use $T = 320$ ps (3.125 GHz), which provides $\sim 7\%$ margin and is a clean round-number value above the search endpoint. The workload is four repetitions of (write all-0s, write all-1s), one cycle per address per pattern, for a total of $16 \times 2 \times 4 = 128$ cycles (≈ 41 ns). The average is taken over the entire run including the initial settling transient, giving $P = 54.74 \mu\text{W}$.

7.3 Figure of Merit

Plugging the measured values $\text{BitcellArea} = 0.308 \mu\text{m}$, $P = 54.74 \mu\text{W}$, and $D = 91.9$ ps into the FOM expression gives

$$\begin{aligned} \text{FOM} &= 60 \cdot 0.308 \mu\text{m} \cdot 54.74 \mu\text{W} \cdot (91.9 \text{ ps})^2 \\ &\approx 8.54 \times 10^{-30} \text{ m} \cdot \text{W} \cdot \text{s}^2. \end{aligned}$$

One caveat is worth flagging: the $D = 91.9$ ps above is measured from the rising edge of ϕ_2 to the point at which the cell has flipped (or the sense-amp output is valid), so it assumes that bitline precharge during the preceding ϕ_1 has already completed. Precharge itself is essential to every access and takes the full preceding half-cycle. If instead D is defined as the time from the start of ϕ_1 (precharge) through to valid data, a more conservative per-operation delay at $T = 320$ ps is $T/2 + 91.9 = 160 + 91.9 = 251.9$ ps, which yields an alternative

$$\begin{aligned} \text{FOM}_{\text{alt}} &= 60 \cdot 0.308 \mu\text{m} \cdot 54.74 \mu\text{W} \cdot (251.9 \text{ ps})^2 \\ &\approx 6.42 \times 10^{-29} \text{ m} \cdot \text{W} \cdot \text{s}^2. \end{aligned}$$

The project specification is ambiguous on which convention is intended, so both numbers are reported.

A straightforward follow-up for reducing the conservative FOM_{alt} would be to upsize the precharge PMOS in the column driver (Section 3.5), which is currently $W = 2$. A wider device would pull the bitlines up faster, shortening ϕ_1 and therefore shrinking the $T/2$ precharge term that dominates D_{alt} . This also raises the maximum operating frequency. I did not have time to sweep this parameter and rerun the binary search, but it is the most obvious lever for pulling FOM_{alt} down toward the primary FOM.

7.4 Summary of Metrics

Table 3: Summary of design metrics.

Metric	Value
Technology	22 nm PTM HP
Supply voltage V_{DD}	0.8 V
Array size	16×4
Bitcell sizing (PD : AX : PU)	88 : 44 : 22 nm (4 : 2 : 1)
BitcellArea (sum of widths)	$0.308 \mu\text{m}$
Minimum clock period T_{min}	298 ps
Maximum clock frequency	3.36 GHz
Measurement period T	320 ps (3.125 GHz)
Worst-case read access time	91.9 ps
Worst-case write access time	90.5 ps
Reported delay D	91.9 ps
Average power P	$54.74 \mu\text{W}$
FOM	$8.54 \times 10^{-30} \text{ m} \cdot \text{W} \cdot \text{s}^2$

8 Conclusion

The 16×4 SRAM meets all functional and performance targets. The two largest design decisions, the two-phase clocking scheme with $PCHb = \overline{\phi_1}$ and the isolated latch-type sense amplifier gated by a ϕ_2 -derived SAE, together ensure a clean separation between precharge and access phases and enable reliable sense-amp resolution on a controlled bitline differential. The remaining margin at T_{min} is spent on the write path, which is the most promising target for further optimization.

References

- [1] *Latch-type sense amplifier*, ResearchGate, figure 4.
https://www.researchgate.net/figure/Latch-type-sense-amplifier_fig4_2977800

I, Lucas Krippendorff, certify that I have complied with the University of Pennsylvania’s Code of Academic Integrity in completing this project.