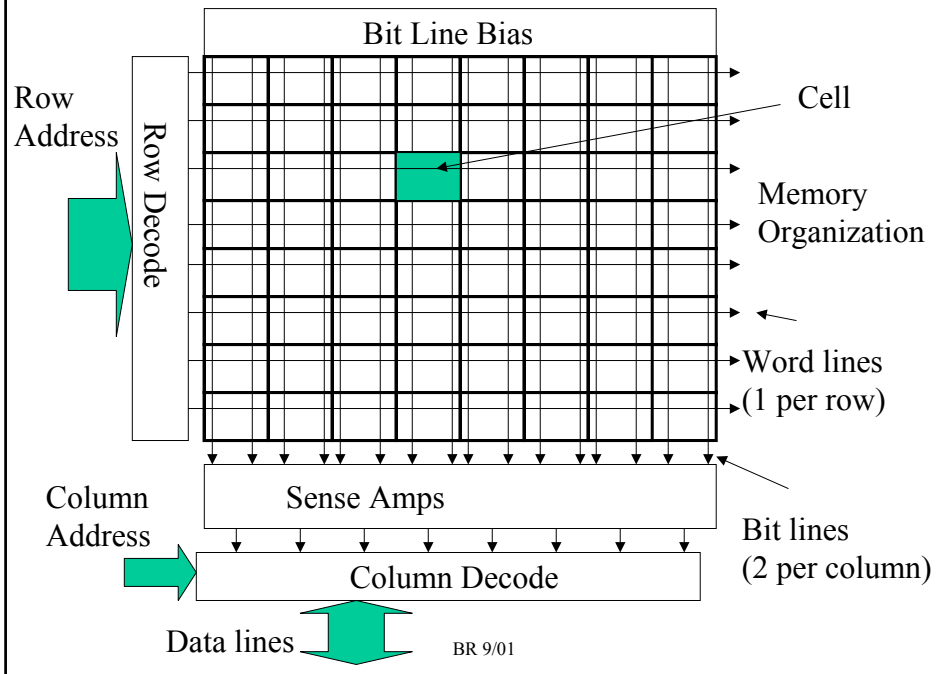


Memory Array

- Typically want an aspect ratio that is not too far from square
- How to divide up the row, column address decoding?

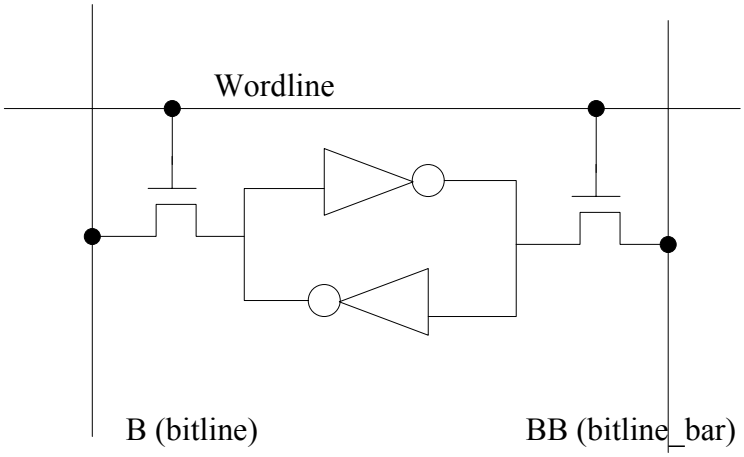
Use an 8K x 32 SRAM = 256 Kb = 2^{18}

2^{18} = 2^9 rows x 2^9 columns

Row decoder is 9 to 512 decoder

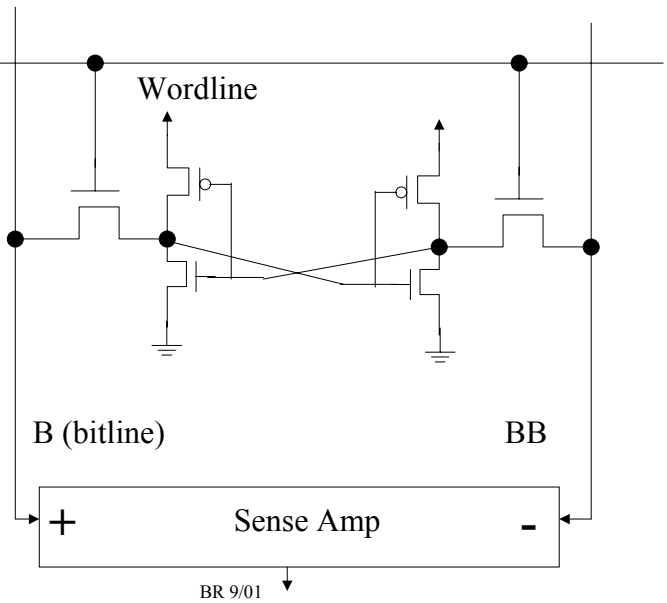
Every 32 (2^5) columns is a 'word', and we only need to decode words. So, column decoder needs to decode 2^4 words, so need a 4 to 16 column decoder.

6-transistor SRAM Cell



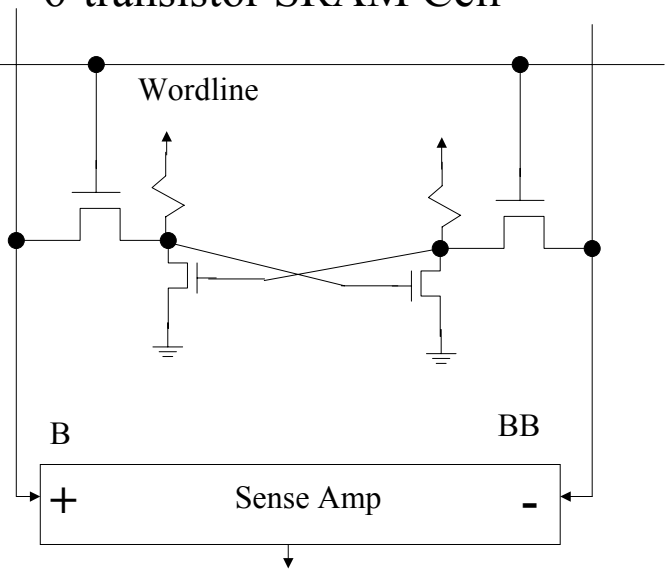
6-transistor SRAM Cell

PMOS loads – no static power dissipation except for leakage, larger cell. Want PMOS small as possible



6-transistor SRAM Cell

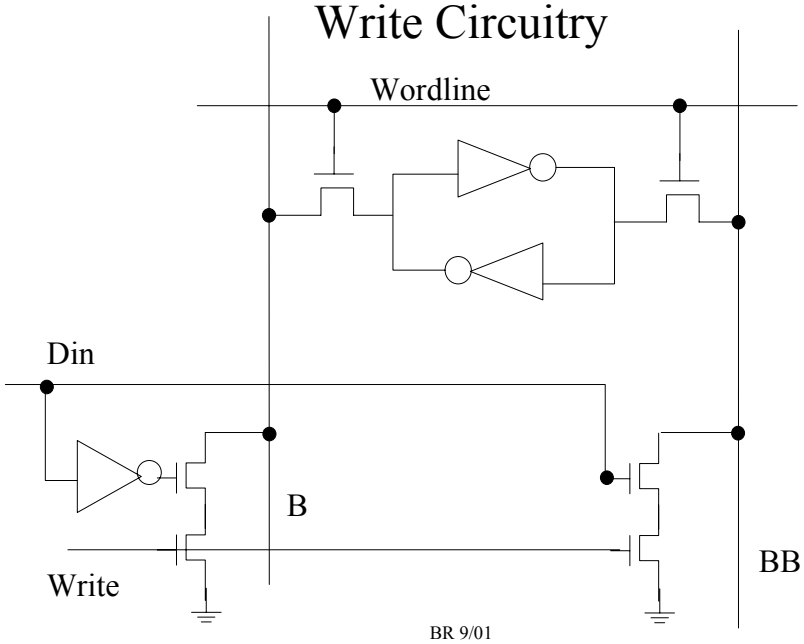
Static loads – smaller cell, faster operation, static power dissipation .



Bit Line vsWord Line

- Assertion of word line accesses all cells in a row
  - Not all bits that are read from a row may be used.
  - Loading on word line is high!
- Bit lines connect all cells in a column, only one cell in a column can ever be on at a time
  - Would like to keep bitline swing low in order to preserve power
  - Sense Amp function is to detect bit line change and produce a full ‘0’ or ‘1’ for output latch

Write Circuitry

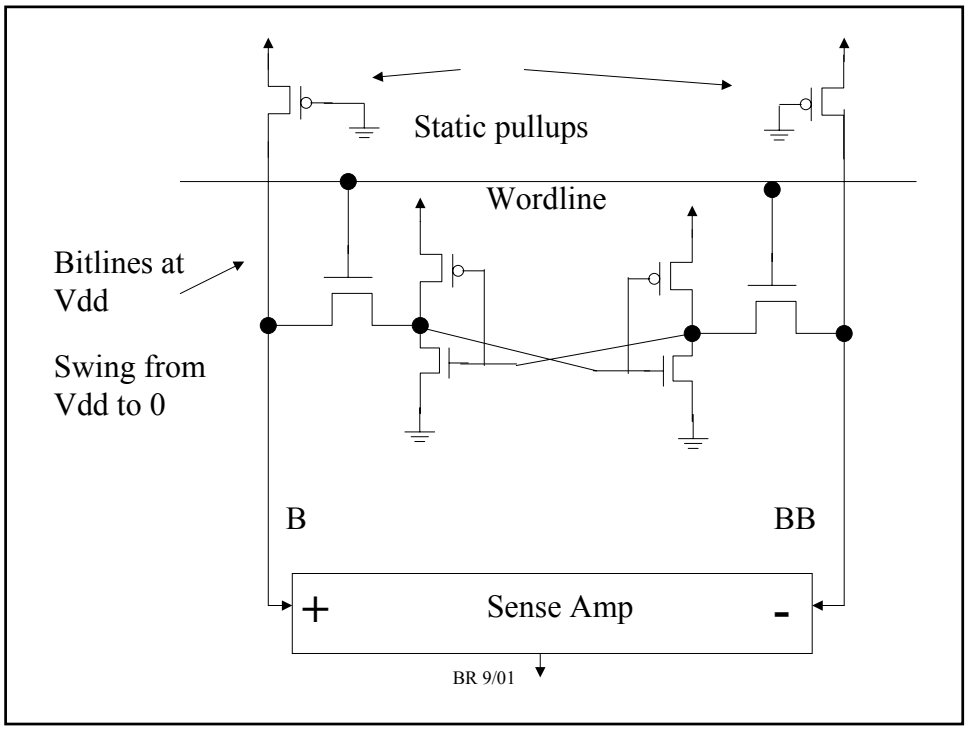


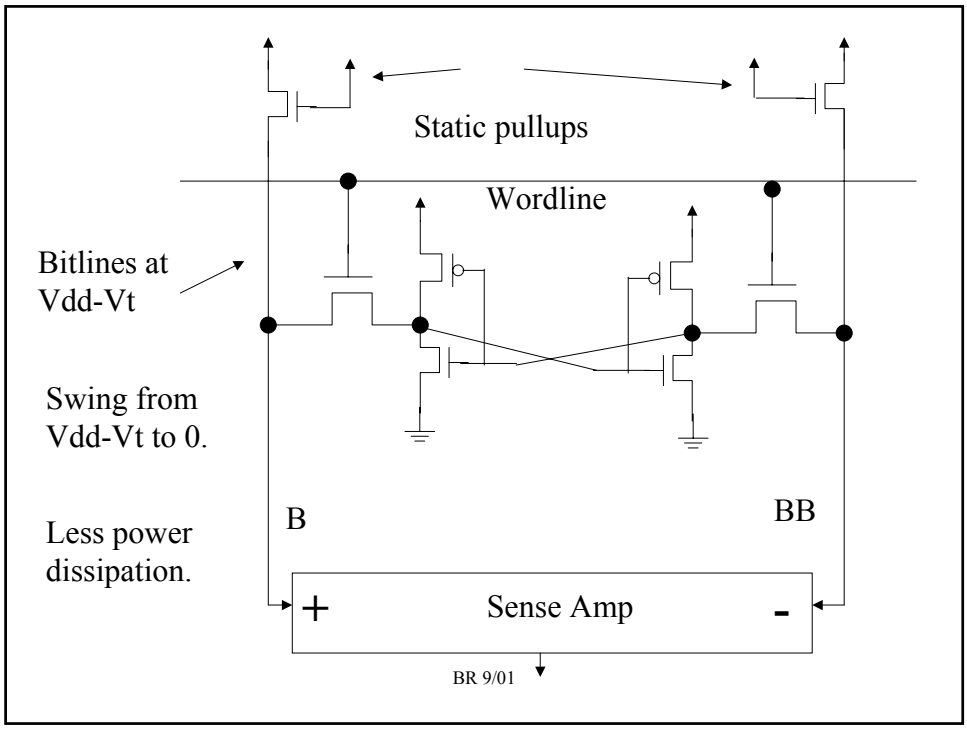
BR 9/01

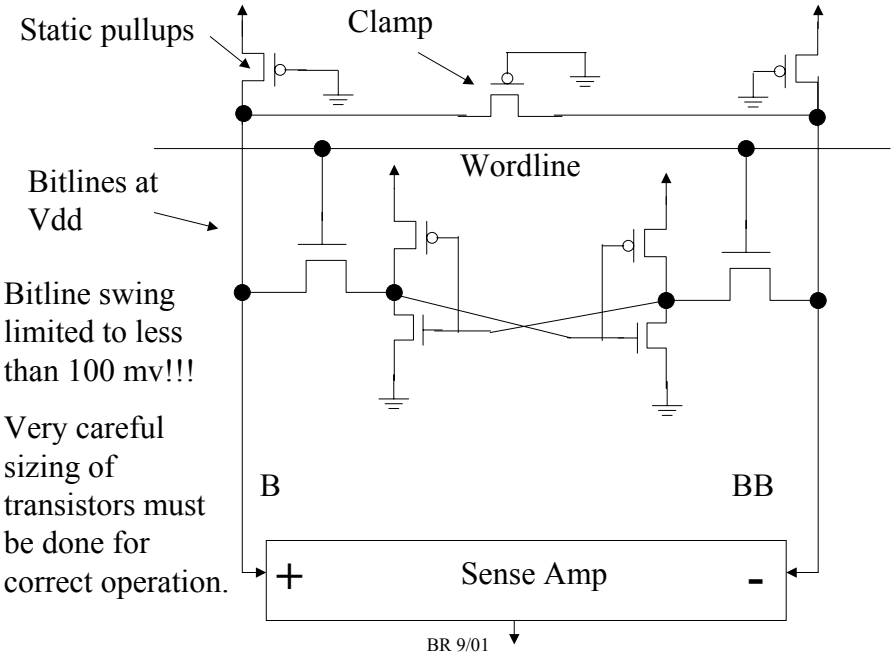
Bit Line Biasing

- Bit lines are typically biased somewhere between 0 and Vdd
  - improve speed of cell for reading/writing
  - limit bitline voltage swing for power savings
- Bit line biasing circuit at top of each column
- Bit line biasing circuit can either be static or dynamic (dynamic operation controlled by clock or pulse generated by address line change)

BR 9/01

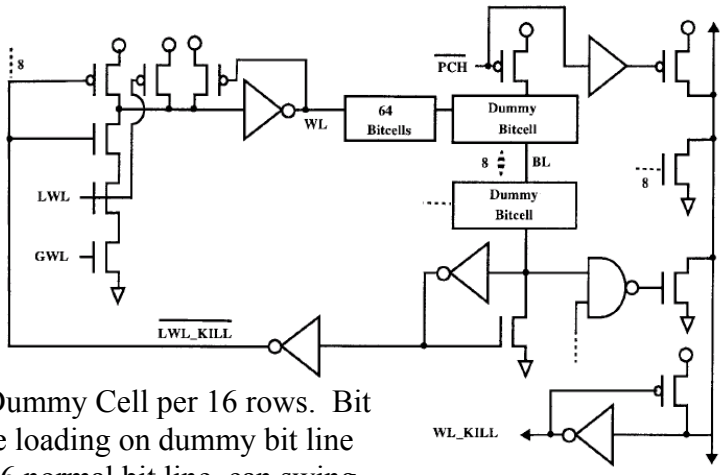

Static pullups
Wordline
Bitlines at
Vdd
Swing from
Vdd to 0
B
BB
+
Sense Amp
-
BR 9/01


Static pullups
Wordline
Bitlines at
Vdd-Vt
Swing from
Vdd-Vt to 0.
Less power
dissipation.
B
BB
+
Sense Amp
-
BR 9/01



BR 9/01

Limiting Bitline Swing via Self-Timed Wordline Kill†



1 Dummy Cell per 16 rows. Bit line loading on dummy bit line 1/16 normal bit line, can swing must faster.

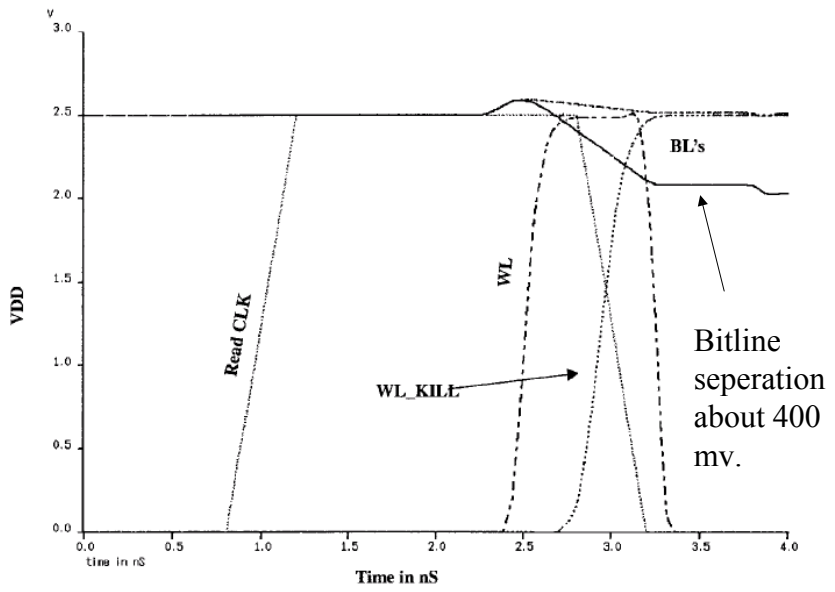
128 rows x 64 columns

BR 9/01

†IJSSC V32, N11, Nov '97, Gerosa, et.al: 250 Mhz, 5-W Powerpc Microprocessor

Word Line Kill Approach

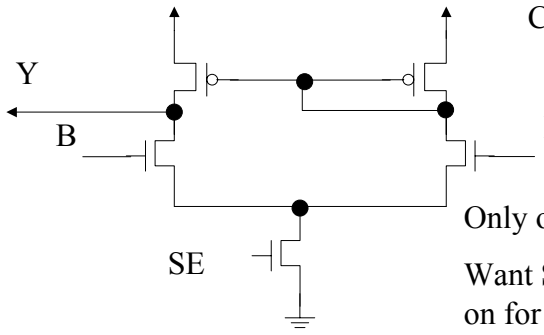
- Only turn on word line for a short period of time
- Will allow RAM cell to only drive bit lines for short period, will limit bit-line swing
- Need self timed circuit to 'turn off' word line
- SRAM easy to self-time by including a dummy cell at END of row (last column is column of dummy cells, 1 dummy cell per 16 rows).
  - When signal arrives at dummy column, then all cells in row have been accessed
- Access of dummy cell feeds back to row decoding to turn OFF word line and turn ON sense amp.



LWL_KILL vs WL_KILL

- Circuit used in L1 Cache of 250 Mhz PowerPC
  - 32 KB divided into 1KB subarrays
  - 1KB subarray 128 rows x 64 columns (8Kb)
- 8 Local Word Line kill lines (1 per 16 rows)
  - Note that LWL_KILL based on bit line output of dummy cell.
  - LWL_KILL turns off decoding for this
- All local word kills are wire or'ed together to form a global word line kill that indicates minimum bit line seperation has occurred, so can turn on sense amp.

Sense Amp (pg 598, Rabaey)



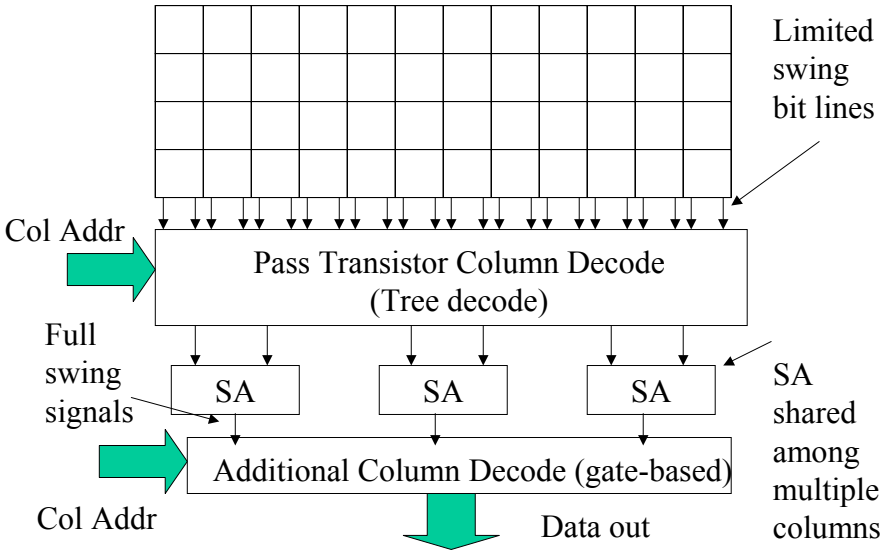
Current-mirror SA

Only one bit line will swing.

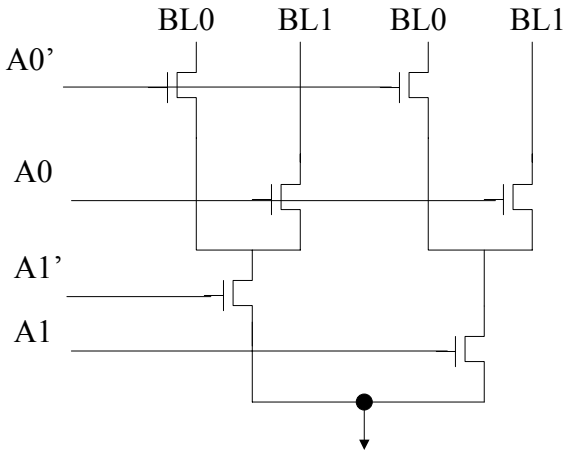
Want Sense Amplifier turned on for short amount of time in order to save power.

Job of SA is to sense bit line swing, amplify to full swing output.

Sharing Sense Ampliers



4 to 1 Tree Decoder (pg. 595, Rabaey)



Are able to use nmos-only pass transistors because of limited swing.

Number of pass transistors in series is a concern, but limited swing helps speed.

Improving Speed, Saving Power

Critical path runs through row decode, word line assertion

- Need smaller decoding, less word line capacitance in order to improve speed.
- Break a large array into smaller sub-arrays, and use hierarchical decoding to select a sub array
  - PowerPC 32K x 8 cache broken into 32 blocks, each 1K x 8
  - Cypress 1Mb Dual Port broken into 32 blocks, each 32 K bits (2^5 x 2^5 x 2^{10} = 2^{20}). Each blocks is 512 rows x 64 columns
  - Mitsubishi SRAM (Rabaey text). 32 blocks of 128K bits (1024 rows x 128 columns)
- Only one sub-array will be activated, saves power!!!!