

The Memory Subsystem



ECE 153B

Sensor & Peripheral Interface Design

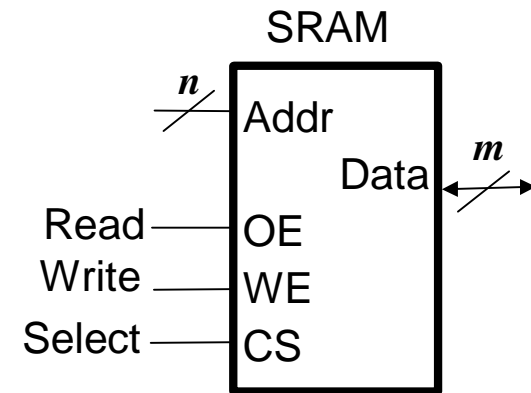
Winter 2016

The Memory Subsystem

- Except for the CPU, the most important subsystem in the computer
 - all von Neumann digital computers have memory
 - modern systems have multiple, independent memories
 - often on-chip ... our 'C31 CPU has two small on-chip SRAMs
- Do not confuse the addressable memory space with the actual memory subsystem
 - number of bits in processor-generated address can be more (or less) than number of bits used to address physical memory
 - some setups have holes in memory, others have MMUs (i.e. memory management units)

SRAM

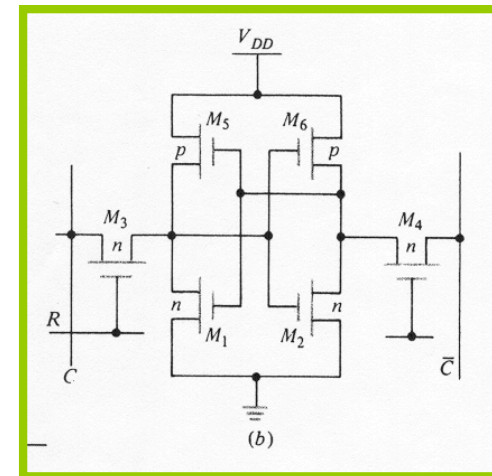
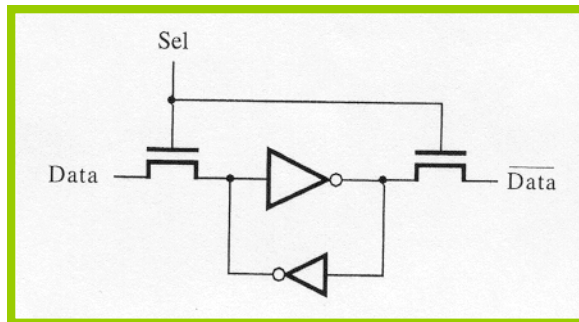
- ❑ Static Random Access Memory
 - essentially a combinational table look-up (also writable)
 - easiest type of memory to use
 - fast, often rather expensive, large high pinout packages
- ❑ Logic symbol needs:
 - 1 wire for requesting a read (aka OE)
 - m wires for the data
 - 1 wire for requesting a write (aka WE)
 - n wires for addressing 2^n cells
 - usually a chip select (1 wire)
- ❑ So-called “separate I/O” variant exists too
 - a second m -wide data path (separate dedicated paths for reading and for writing)



SRAM Cell

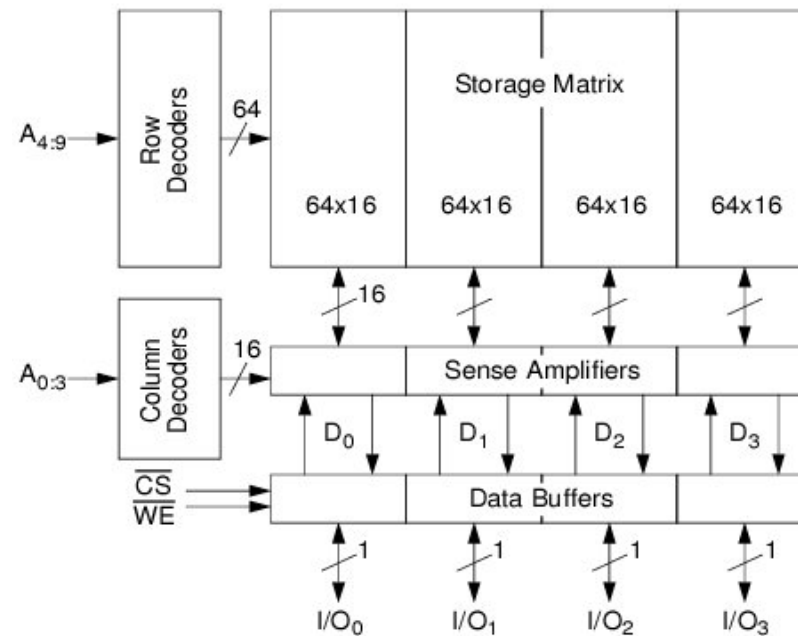
□ 6T SRAM Cell

- CMOS implementation with pass transistors
- Sense amp at bottom of column



SRAM Structure

- SRAM internal block diagram
 - e.g. 1024 x 4-bit SRAM



- sense amplifiers on bit and $\bar{\text{bit}}$ — detect value on read
- data pins often shared as inputs and outputs
- write by forcing bit and $\bar{\text{bit}}$ to appropriate values

SRAM Uses & Properties

- SRAM is typically used for
 - register files
 - caches
 - processor on-chip memories

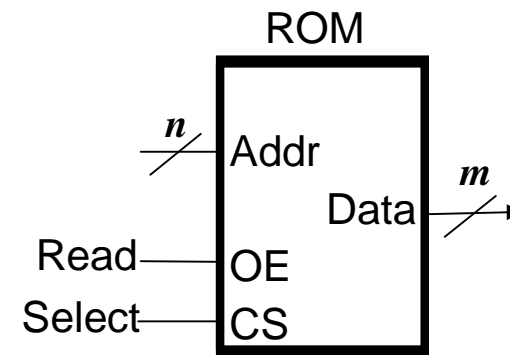
- Capacity (relative to denser DRAMs)
 - about $\frac{1}{4}$ or less
 - price higher, due primarily to larger die size & larger packaging

- Speed (relative to DRAMs)
 - Faster; SRAM is the fastest memory available

- Power (for CMOS): static \approx zero
 - Dynamic power proportional to usage

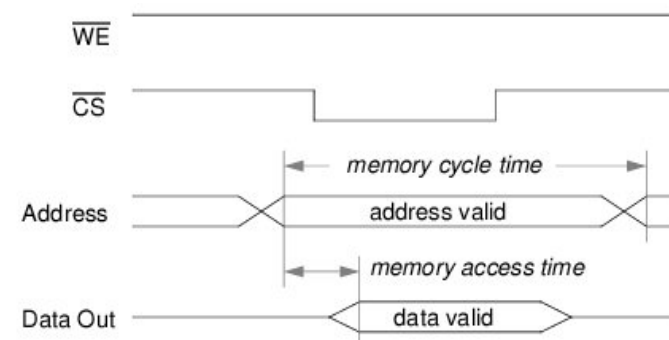
ROM

- Read Only Memory
- ROM, PROM, EPROM, EEPROM....
 - not in-system writeable, except by special means
- Logic symbol has:
 - 1 wire for requesting a read (aka OE)
 - m wires for the data
 - n wires for addressing 2^n cells
 - often a chip select (1 wire)
- Uses:
 - non-volatile information (usually parameters)
 - initial program load (a.k.a. IPL or boot-up)
- Speed (often quite slow)
- Cost (cheap)



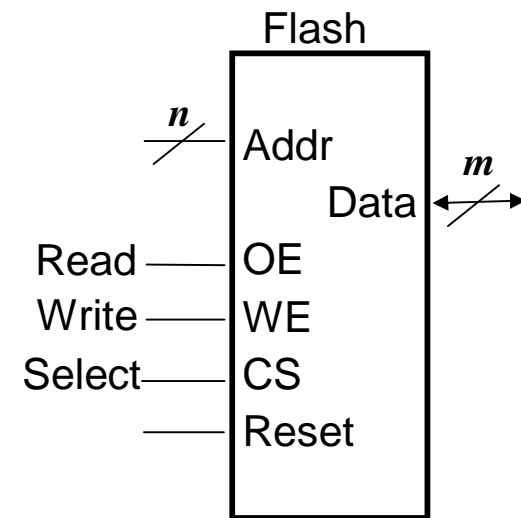
SRAM, ROM & Flash Memory Cycles

- With SRAMs, ROMs & Flash
 - no notion of a “cycle”
 - asynchronous interface
 - latency (called the “access time”)
- To read:
 - present the address
 - assert OE (output enable) and CS
 - wait for access time, t_{acc}
 - latch the data
- To write into an SRAM (doesn't apply to ROM or Flash):
 - present the address and the data
 - assert WE (write enable) and CS
 - deassert WE before changing either address or data



Flash Memory (& Read Mostly)

- Flash, battery-backed SRAM, various floating gate SRAMs
 - all are in-system writable, but write access isn't their primary usage
- Logic symbol needs:
 - 1 wire for requesting a read (aka OE)
 - 1 wire for requesting a write (aka WE)
 - m wires for the data
 - n wires for addressing 2^n cells
 - often a wire for (flash) FSM reset
 - often a chip select (1 wire)
- Uses:
 - non-volatile information (usually parameters)
 - initial program load (a.k.a. IPL or boot-up)
- Speed (fast for read, very slow & complex for writing)
- Cost (relatively inexpensive)



Flash Memory Characteristics

- Flash memories are non-volatile
 - Often block-oriented, usually not single-word writeable
 - Erase of an entire block must be done first ... before any writing
 - Blocks can be uniform or irregular in size (per flash & manufacturer)
 - Once erased and written, data is retained for 50+ years

Writing to Flash Memory (a.k.a., Programming)

- ❑ To write into a Flash (varies by manufacturer)
 - Ensure internal state machine has been reset
 - Perform flash's "erase sequence"
 - ❑ usually 6 SRAM-like writes with a specific pattern of addresses and data
 - Perform "write unlock" sequence
 - ❑ 4-6 writes to specific pattern of addresses and data
 - Perform a single-word write (with your data), just like an SRAM
 - Repeat the above "unlock" and "write" for each word (or block)
 - Must wait for each write to complete (by polling or other timing)

NOR vs. NAND Flash

- NOR and NAND flash differ in two important ways
 - The connections of the individual memory cells are different
 - In NOR flash, cells are connected in parallel to the bit lines which resembles the parallel connection of transistors in a CMOS NOR gate
 - In NAND flash, cells are connected in series, resembling a NAND gate
 - The interface provided for reading and writing the memory is different
 - NOR allows random-access for reading
 - NAND allows only page access

NOR vs. NAND Flash

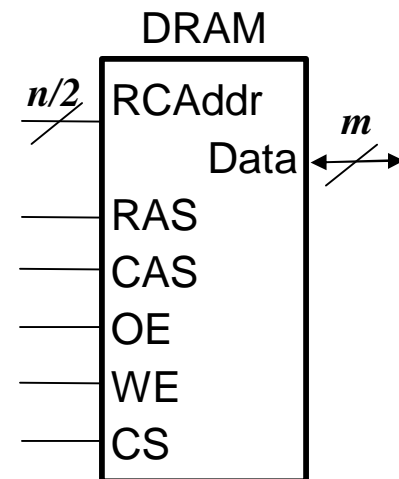
- ❑ NOR offers faster read speed and random access capabilities
 - Suitable for code storage in devices such as PDAs and cell phones
- ❑ NOR write and erase functions are slow compared to NAND
- ❑ NOR has a larger memory cell size than NAND
 - Limits scaling capabilities and therefore achievable bit density compared to NAND
 - Since code storage tends to require lower density memory than file storage, NOR's larger cell size is not considered a concern when used in these applications

NOR vs. NAND Flash

- NAND offers fast write/erase capability and is slower than NOR in the area of read speed
 - NAND is, however, more than sufficient for a majority of consumer applications such as digital video, music or data storage
- NAND's fast write/erase speed combined with its
 - Higher available densities and a
 - Lower cost-per-bit than NOR
 - make it the favored technology for file storage in a host of consumer applications
 - NAND is typically used for
 - storing large quantities of information in devices such as
 - flash drives, MP3 players, multi-function cell phones, digital cameras and USB drives

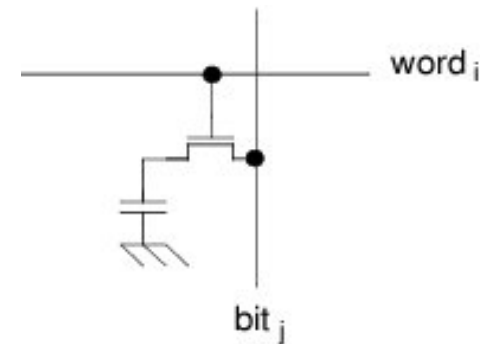
DRAM

- Dynamic Random Access Memory
 - data stored as charge on a tiny capacitor (MOSFET-gate)
 - organized in rows and columns, addressed separately & sequentially
 - not the same as SDRAM (synchronous DRAM)
- Logic symbol needs:
 - 1 wire for enabling data outputs (aka OE), often omitted
 - 1 wire for requesting a write (aka WE)
 - m wires for the data
 - $n/2$ wires for addressing $2n$ cells (second generation)
 - 2 wires for row vs column address select
 - often a chip select (1 wire)
- Capacity: high
- Speed (fast, but not as fast as SRAM due to more complex cycle)
- Cost (relatively inexpensive)



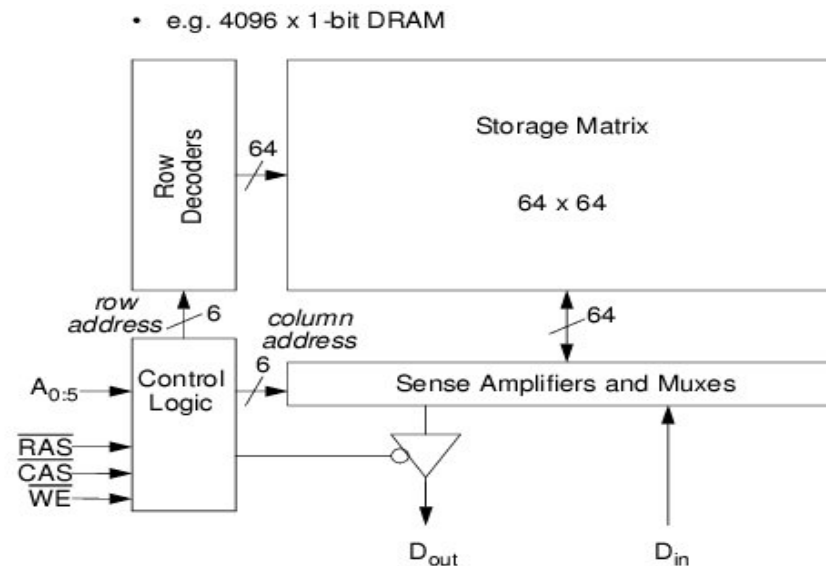
DRAM Refresh Cycles

- Improved density & lower cost (relative to SRAMs) due to
 - use of stored charge (1 transistor) vs. much more complex circuit
 - use of fewer pins (thus smaller package) for row/column addressing
- DRAMs internally organized as a square array
 - \sqrt{n} rows by \sqrt{n} columns for n cell DRAM
- Memory must be “refreshed” periodically
 - by reading or writing at least once to each row
 - refresh period is in the range 16-32 milliseconds
- Example (considering a 64M x 8 part)
 - $\sqrt{64M} = 8K$ rows
 - 8K refreshes per 32 msec \implies 1 row refreshed every 4 μ sec



DRAM Addressing

- One-half of the address pins saved (relative to SRAM)
 - By specifying row and column addresses sequentially
 - Requires separate row_address_strobe (RAS) & column_addr_strobe (CAS)

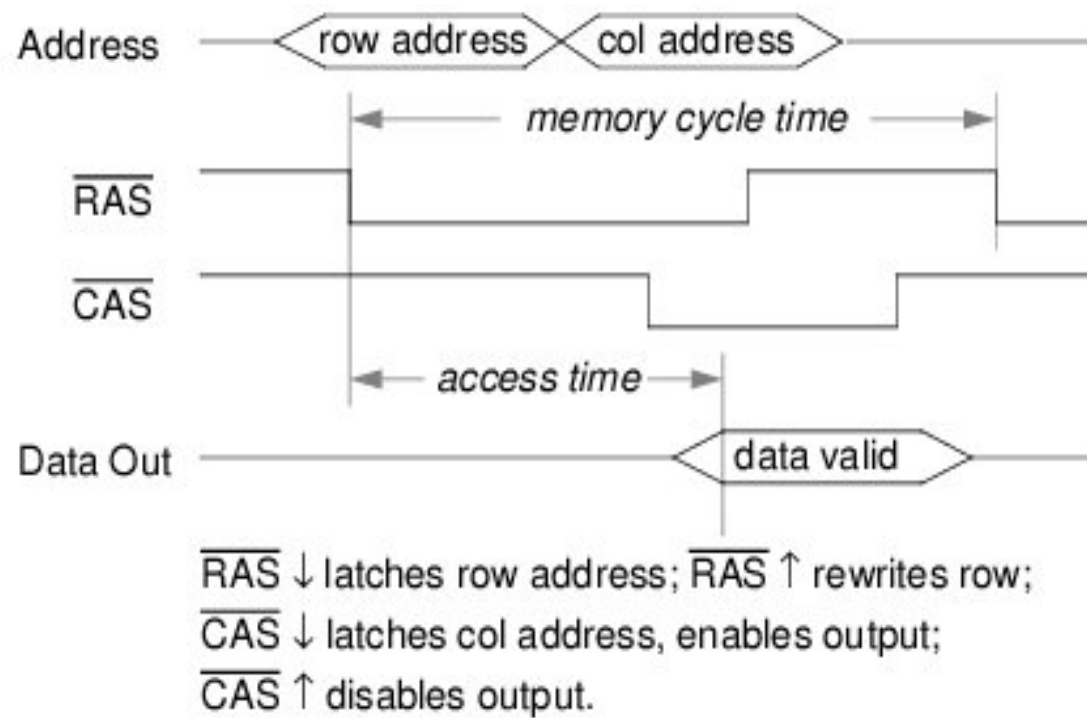


DRAM Addressing

- ❑ Second generation DRAM timing and sequencing are complex
 - Cycle types:
 - ❑ random read, random write
 - ❑ fast page mode read, fast page mode write
 - ❑ CBR refresh (CAS before RAS)
 - ❑ “hidden” refresh
 - ❑ read/modify/write

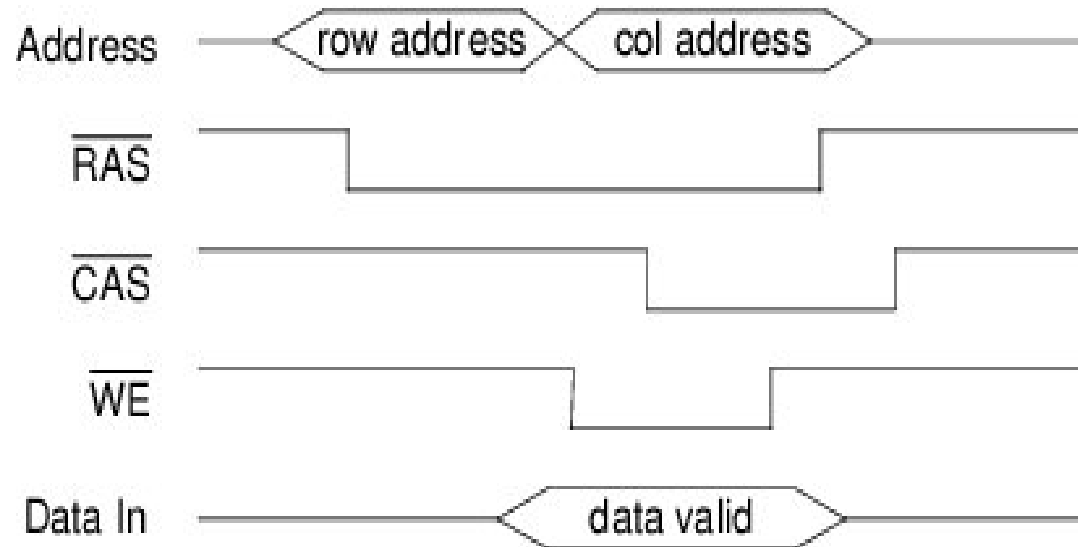
- ❑ Other variants exist as well:
 - nibble mode reads & writes
 - static column
 - RAS-only refresh

DRAM Random Read Cycle



DRAM Random Write Cycle

- $\overline{\text{RAS}}/\overline{\text{CAS}}$ sequencing for write

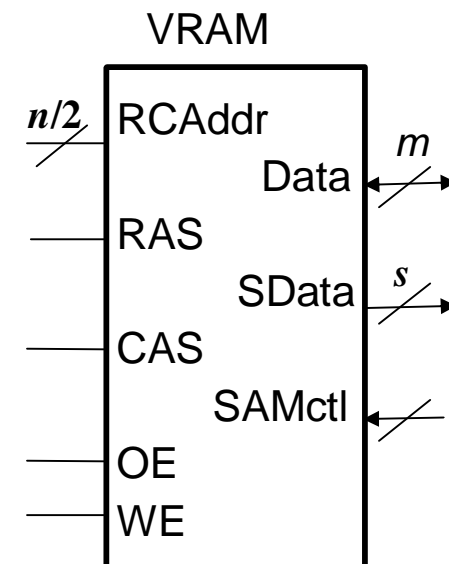


High-Performance DRAM Cycles

- ❑ Optimizations to get more bits per RAS
- ❑ Nibble mode (1 nibble == 4 sequential bits or cells)
 - present RAS, then CAS, CAS, CAS, CAS
- ❑ Fast page mode
 - present RAS, then provide col addr, then CAS
 - then change to different column addr (within same row), then CAS again ...
- ❑ Static column
 - present RAS, then column address and CAS
 - then change column addr (like an SRAM) without cycling CAS

VRAM

- Video RAM
 - specialty DRAM optimized for graphics applications
 - regular DRAM with an added sequential serial read port
- Logic symbol for regular DRAM, plus:
 - s wires (e.g. 1,4,8) for serial data
 - 1-2 wires for serial controls (SAMctl)
- Capacity: same as DRAM
- Speed (same as DRAM)
- Cost (slightly more than DRAM)
- Complex interactions possible between parallel and serial sides since the internal memory is shared
- Replaced by SDRAM (next topic) in most modern graphics applications



SDRAM

- Synchronous DRAM
 - Third generation DRAM
- Made by adding a synchronous interface between the basic core DRAM operation/circuitry of second-generation DRAMs and the control coming from off-chip to make the DRAM operation faster
 - All commands and operations to and from the DRAM are executed on the rising edge of a master or command clock signal
- Additionally, the memory is segmented into banks
 - The bank selected is determined by the addresses BA1 and BA0

SDRAM

- Advantages of banked SDRAM
 - Possible to activate a row in one bank and then, while the row is opening, perform an operation in some other bank (such as reading or writing)
 - In addition, one of the banks can be in a PRECHARGE mode (the bitlines are driven to $V_{cc}/2$) while accessing one of the other banks and, thus, in effect hiding PRECHARGE and allowing data to be continuously written to or read from the SDRAM

SDRAM

- SDRAMs often employ pipelining in the address and data paths to increase operating speed
 - Disconnects operating frequency from access latency
 - Additional access instructions can be fed into the SDRAM before prior access instructions have completed

DDR SDRAM

- ❑ Double Data Rate Synchronous Dynamic Random Access Memory
- ❑ Compared to single data rate (SDR) SDRAM, the DDR SDRAM interface makes higher transfer rates possible by more strict control of the timing of the electrical data and clock signals
- ❑ Implementations often have to use schemes such as phase-locked loops and self-calibration to reach the required timing accuracy
- ❑ The interface uses double pumping (transferring data on both the rising and falling edges of the clock signal) to lower the clock frequency
 - One advantage of keeping the clock frequency down is that it reduces the signal integrity requirements on the circuit board connecting the memory to the controller
- ❑ The name "double data rate" refers to the fact that a DDR SDRAM with a certain clock frequency achieves nearly twice the bandwidth of a SDR SDRAM running at the same clock frequency, due to this double pumping
- ❑ DDR SDRAM, also called DDR1 SDRAM, has been superseded by DDR2 SDRAM and DDR3 SDRAM
 - Neither of its successors are forward or backward compatible with DDR1 SDRAM, meaning DDR2 or DDR3 memory modules will not work in DDR1-equipped motherboards, and vice versa.

NXP LPC 4088 Internal Memory & External Memory Controller (EMC)



ECE 153B

Sensor and Peripheral Interface Design

Winter 2015

NXP LPC 4088 Internal Memory

- 512 KB on-chip Flash program memory with In-System Programming (ISP) and In-Application Programming (IAP) capabilities
 - The combination of an enhanced Flash memory accelerator and location of the Flash memory on the CPU local code/data bus provides high code performance from Flash

- 96 KB on-chip SRAM includes 64 kB of main SRAM on the CPU with local code/data bus for high-performance CPU access
 - Two 16 kB peripheral SRAM blocks with separate access paths for higher throughput
 - These SRAM blocks may be used for DMA memory as well as for general purpose instruction and data storage

NXP LPC 4088 Internal Memory

- 4 KB (4032 byte) EEPROM
 - The EEPROM is indirectly accessed through address and data registers
 - All communication with the actual EEPROM block is done through the 64 byte page buffer

- 8 KB Boot ROM
 - The boot loader controls initial operation after reset and also provides the tools for programming the flash memory
 - This could be initial programming of a blank device, erasure and re-programming of a previously programmed device, or programming of the flash memory by the application program in a running system

NXP LPC 4088 Internal Memory

Table 4. LPC408x/7x memory usage and details

Address range	General Use	Address range details and description	
0x0000 0000 to 0x1FFF FFFF	On-chip non-volatile memory	0x0000 0000 to 0x0007 FFFF	For devices with 512 kB of flash memory.
		0x0000 0000 to 0x0003 FFFF	For devices with 256 kB of flash memory.
		0x0000 0000 to 0x0001 FFFF	For devices with 128 kB of flash memory.
		0x0000 0000 to 0x0000 FFFF	For devices with 64 kB of flash memory.
	On-chip SRAM	0x1000 0000 to 0x1000 FFFF	For devices with 64 kB of main SRAM.
		0x1000 0000 to 0x1000 7FFF	For devices with 32 kB of main SRAM.
		0x1000 0000 to 0x1000 3FFF	For devices with 16 kB of main SRAM.
Boot ROM	0x1FFF 0000 to 0x1FFF 1FFF	8 kB Boot ROM with flash services.	
0x2000 0000 to 0x3FFF FFFF	On-chip SRAM (typically used for peripheral data)	0x2000 0000 to 0x2000 1FFF	Peripheral SRAM - bank 0 (first 8 kB)
		0x2000 2000 to 0x2000 3FFF	Peripheral SRAM - bank 0 (second 8 kB)
		0x2000 4000 to 0x2000 7FFF	Peripheral SRAM - bank 1 (16 kB)
	AHB peripherals	0x2008 0000 to 0x200B FFFF	See Figure 9 for details
0x4000 0000 to 0x7FFF FFFF	APB Peripherals	0x4000 0000 to 0x4007 FFFF	APB0 Peripherals, up to 32 peripheral blocks of 16 kB each.
		0x4008 0000 to 0x400F FFFF	APB1 Peripherals, up to 32 peripheral blocks of 16 kB each.

NXP LPC 4088

External Memory Controller (EMC)

- ❑ Dynamic memory interface support including single data rate SDRAM
- ❑ Asynchronous static memory device support including RAM, ROM, and Flash, with or without asynchronous page mode
- ❑ Low transaction latency
- ❑ Read and write buffers to reduce latency and to improve performance
- ❑ 8/16/32 data and 16/20/26 address lines wide static memory support
- ❑ 16 bit and 32 bit wide chip select SDRAM memory support
- ❑ Static memory features include:
 - Asynchronous page mode read
 - Programmable Wait States
 - Bus turnaround delay
 - Output enable and write enable delays
 - Extended wait

NXP LPC 4088

External Memory Controller (EMC)

- ❑ Four chip selects for synchronous memory and four chip selects for static memory devices
- ❑ Power-saving modes dynamically control EMC_CKE and EMC_CLK outputs to SDRAMs
- ❑ Dynamic memory self-refresh mode controlled by software
- ❑ Controller supports 2048 (A0 to A10), 4096 (A0 to A11), and 8192 (A0 to A12) row address synchronous memory parts
 - That is typical 512 MB, 256 MB, and 128 MB parts, with 4, 8, 16, or 32 data bits per device
- ❑ Separate reset domains allow the for auto-refresh through a chip reset if desired.
- ❑ Synchronous static memory devices (synchronous burst mode) are not supported.

NXP LPC 4088

External Memory Controller (EMC)

Table 4. LPC408x/7x memory usage and details

Address range	General Use	Address range details and description	
0x8000 0000 to 0xDFFF FFFF	Off-chip Memory via the External Memory Controller	Four static memory chip selects:	
		0x8000 0000 to 0x83FF FFFF	Static memory chip select 0 (up to 64 MB)
		0x9000 0000 to 0x93FF FFFF	Static memory chip select 1 (up to 64 MB)
		0x9800 0000 to 0x9BFF FFFF	Static memory chip select 2 (up to 64 MB)
		0x9C00 0000 to 0x9FFF FFFF	Static memory chip select 3 (up to 64 MB)
		Four dynamic memory chip selects:	
		0xA000 0000 to 0xAFFF FFFF	Dynamic memory chip select 0 (up to 256 MB)
		0xB000 0000 to 0xBFFF FFFF	Dynamic memory chip select 1 (up to 256 MB)
		0xC000 0000 to 0xCFFF FFFF	Dynamic memory chip select 2 (up to 256 MB)
		0xD000 0000 to 0xDFFF FFFF	Dynamic memory chip select 3 (up to 256 MB)
0xE000 0000 to 0xE00F FFFF	Cortex-M4 Private Peripheral Bus	0xE000 0000 to 0xE00F FFFF	Cortex-M4 related functions, includes the NVIC and System Tick Timer.

External Memory Controller Block Diagram

