A History and Future of Memory Innovation

Dean A. Klein

VP Advanced Memory Solutions

Micron Technology, Inc.

SEMICON® CHINA

Micron®

The Role of Memory

"The necessity of an inner memory in electronic digital computers has been realized by all designers. The high computing speed possible with electronic devices becomes useful only when sufficient intermediary results can be memorized rapidly to allow the automatic handling of long sequences of accurate computations which would be impractically lengthy by any other slower means. An ideal inner memory organ for a digital computer should be able to register in as short a writing time as possible any selected one of as many as possible on-off signals and be able to deliver unequivocally the result of this registration after an arbitrarily long or short storing time with the smallest possible delay following the reading call." – Jan Rajchman

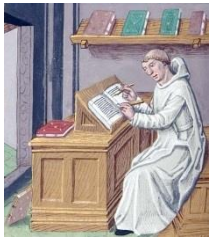
Early Memory

Papyrus

Paper

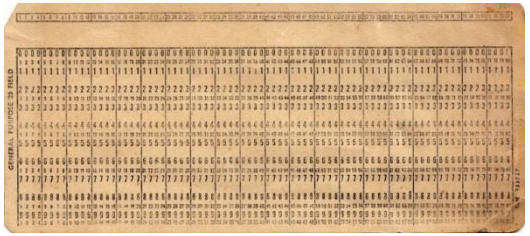
Paper tape

Parchment



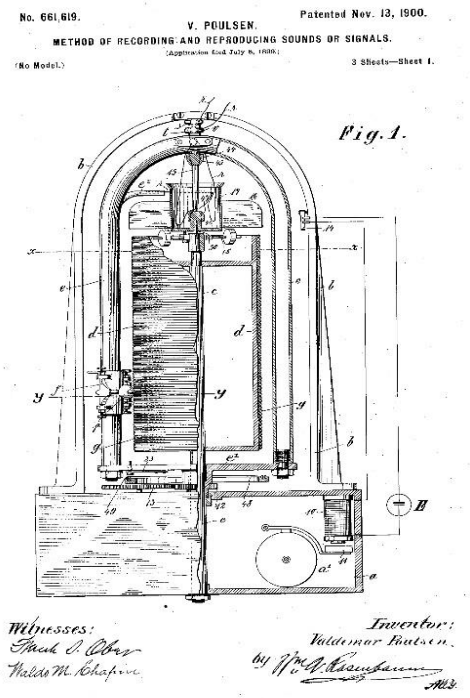
Vellum

Punch card (“H-Card”)



Magnetic Wire

- Valdimar Poulsen's magnetic wire patent 1899


No. 661,619. Patented Nov. 13, 1900.

V. POULSEN.

METHOD OF RECORDING AND REPRODUCING SOUNDS OR SIGNALS.

(No Model.) 3 Sheets—Sheet 1.

Fig. 1.

Witnesses:
Frank L. Ober
Waldo M. Chapin

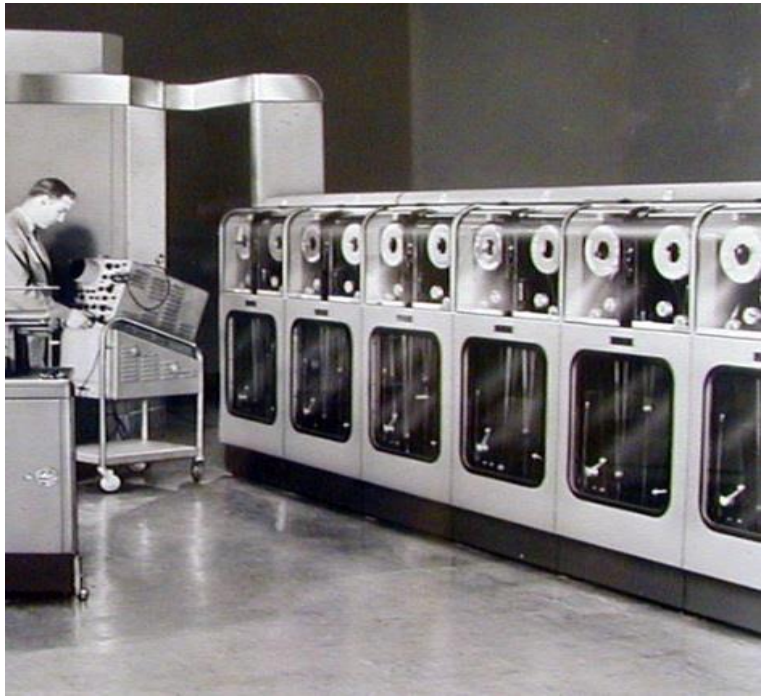
Inventor:
Valdemar Poulsen.
by Wm. W. Dyer & Son
Atty.

Magnetic Tape

- Austrian Fritz Pfleumer
- Paper tape with iron oxide
- Used for audio

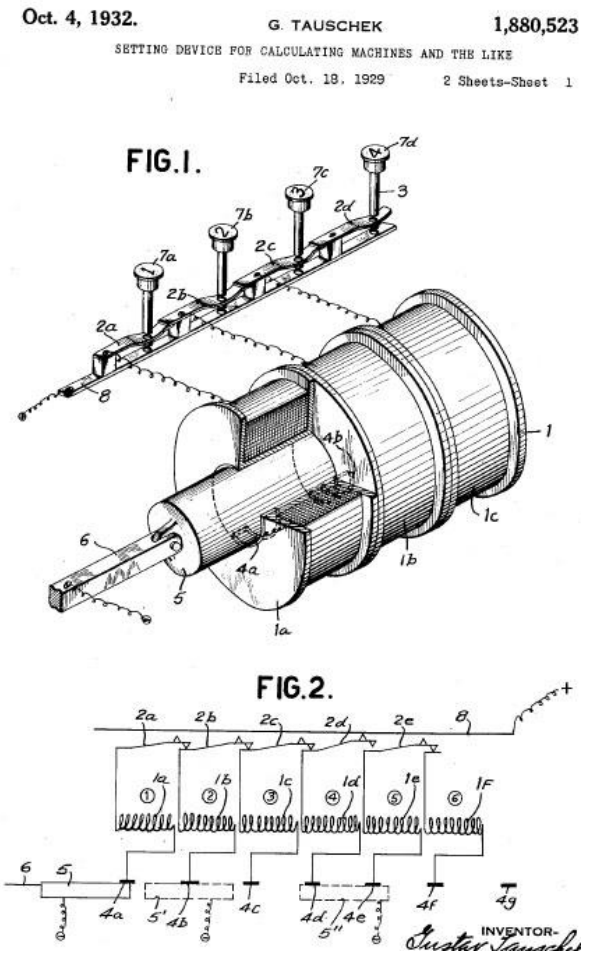
Univac-1 Magnetic Tape - 1951

- 128 characters per inch – 25.6K characters/in^2
- 12,800 characters/sec, 7,200 usable



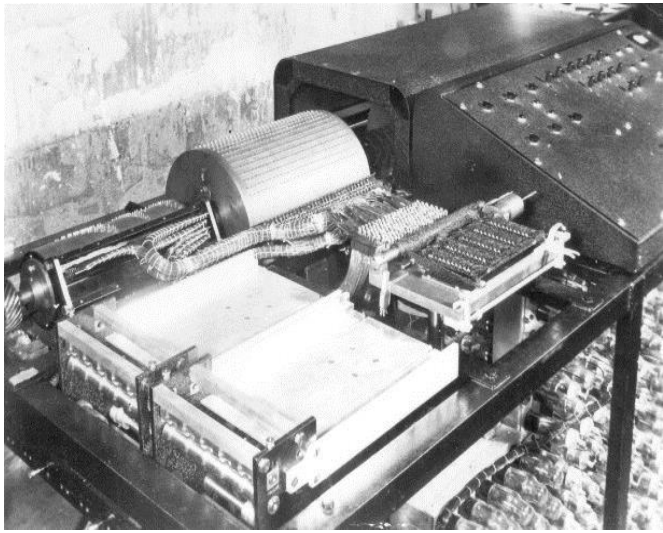
Drum Memory

- Magnetic drum memory was invented by Gustav Tauschek in 1932 in Austria



Capacitive Drum Memory – The ABC

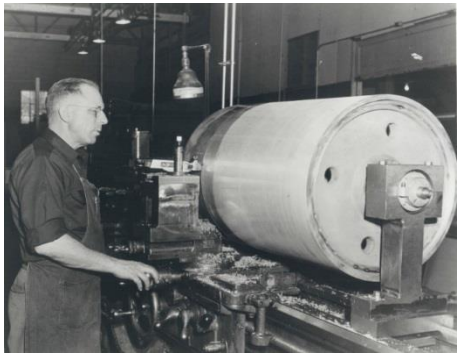
- 1941 the Atanasoff-Berry Computer.
- Called a "regenerative capacitor memory," the system used a pair of drums, each containing 1600 capacitors, with connections to the capacitors covering the surface of the drum and rotating at one revolution per second.
- The system gave 3000 bits of total usable memory.
- The core of the computational engine was synchronized directly to the drum.
- Similar to another memory type that would make its appearance some 25 years later, regenerative capacitor memory required periodic refresh…



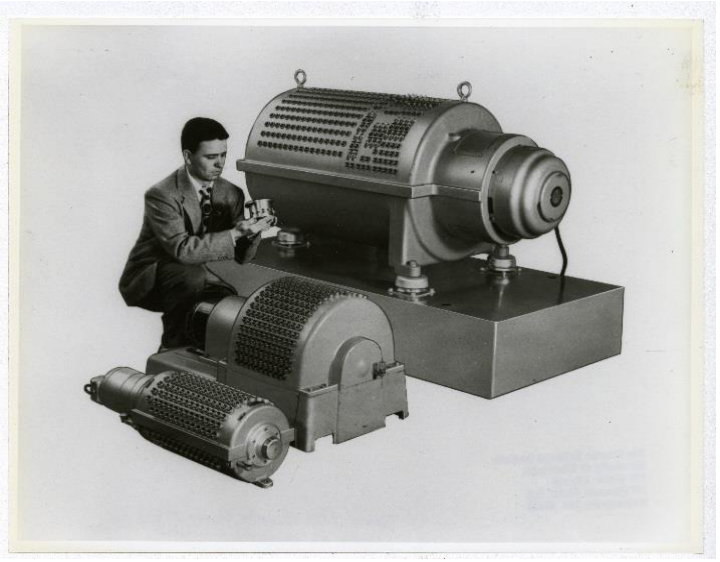
Atanasoff-Berry computer. Courtesy University of MN, Charles Babbage Institute

The UNIVAC 1101 (aka ERA 1101) - 1948

- Used drum memory for main memory.
- Drum diameter of 8.5 inches.
- Spinning speed of 3500RPM.
- 200 R/W heads, 48KB capacity.
- Access time of 32µsec to 17msec



An ERA machinist making a magnetic drum. Courtesy University of MN, Charles Babbage Institute



ERA magnetic drum memories. Donald Weidenbach pictured. Courtesy University of MN, Charles Babbage Institute

Other Notable Drum Memory Machines

Atlas operator console

IBM 650 commercial computer

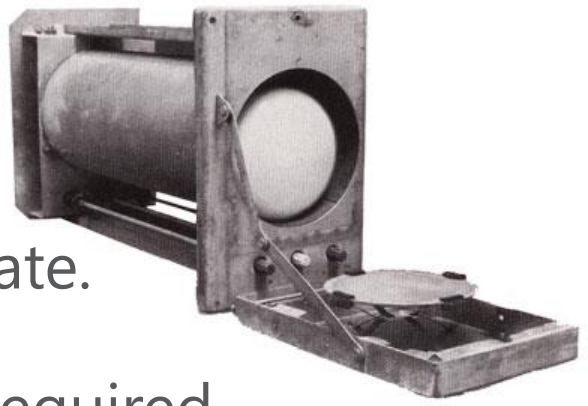
- In 1958 the British Atlas supercomputer was built around a four-drum 576KB memory system.
- The first commercial mass-produced computer, the IBM 650, had an 8.5KB drum memory rotating at 12,500RPM.

1940's Sidetrack: Delay Line Memories

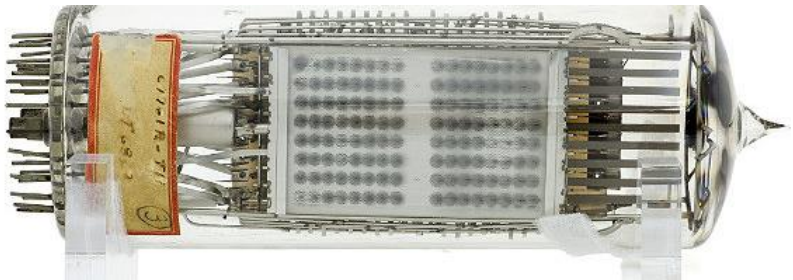
- Mercury-filled glass tubes and quartz transducers plugging each end
- Stored information as sound waves in the mercury.
- These memories were eventually able to store up to a few thousand bits per tube.
- 1949 British EDSAC computer used 32 delay lines of 576 bits each.
- UNIVAC 1 used 18 120-bit mercury delay line columns with an average access time of just over 200µsec

EDSAC mercury delay lines. M.V. Wilkes pictured. Courtesy of the Computer Lab at Cambridge University

Williams Tubes and Selectron Tubes

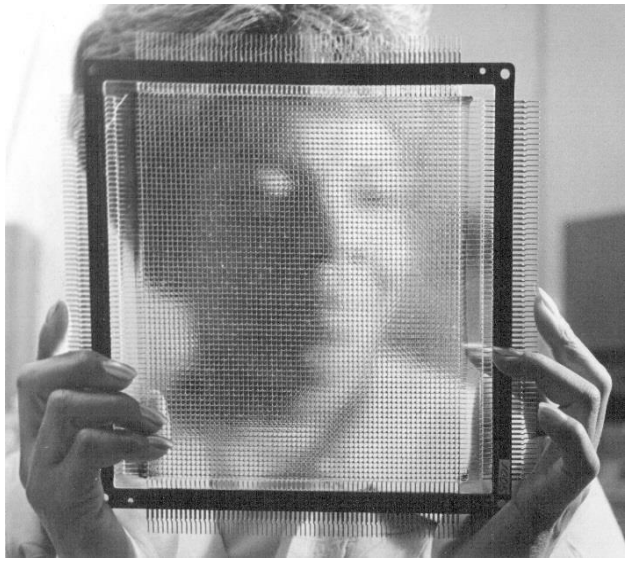


- Used a CRT to create charge wells on the face of the tube, sensed by a read plate.
- Williams tube: Destructive read, refresh required, random access.
  - 2K-4K bits
- Selectron tube: Non-destructive read, longer refresh, random access.
  - 256 bits



Core Memory

- Developed and commercialized in the 1947 to 1952
- First used in the MIT Whirlwind computer.
- Initial core density of 2K bytes, installed in the Whirlwind in 1953.
- Over its life, core memory costs went from $1/bit to $0.01/bit.
- Adoption almost ubiquitous.



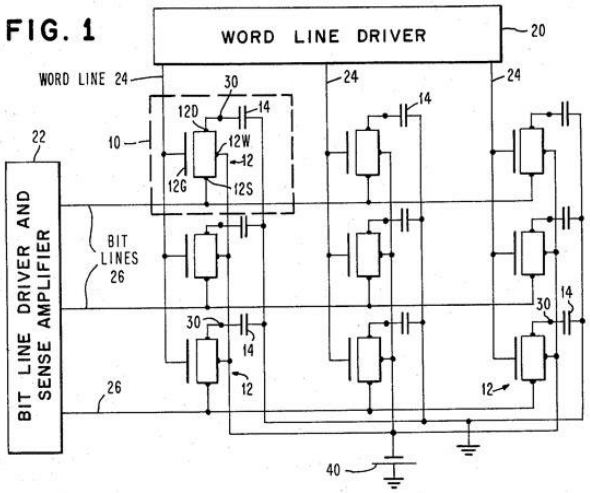
Magnetic core memory. Courtesy University of MN, Charles Babbage Institute

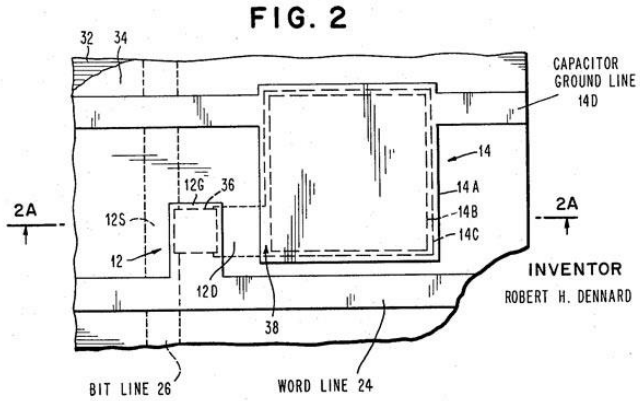
Machine Evolution

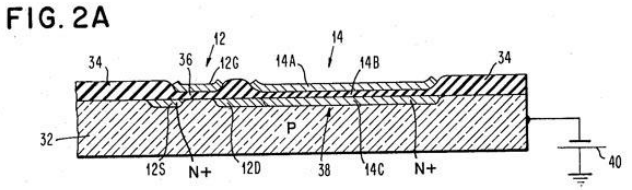
- External “storage” vs. internal “memory”.
- Hard disk technology produced in 1956:
  - IBM 350 RAMAC
  - 50-disk, one-ton
  - Capacity of 3.75 megabytes.
  - Average access time of about 600μsec
  - Areal density of about 2000 bits per square inch
  - Cost of about $9000 per megabyte.

Something New: Semiconductors

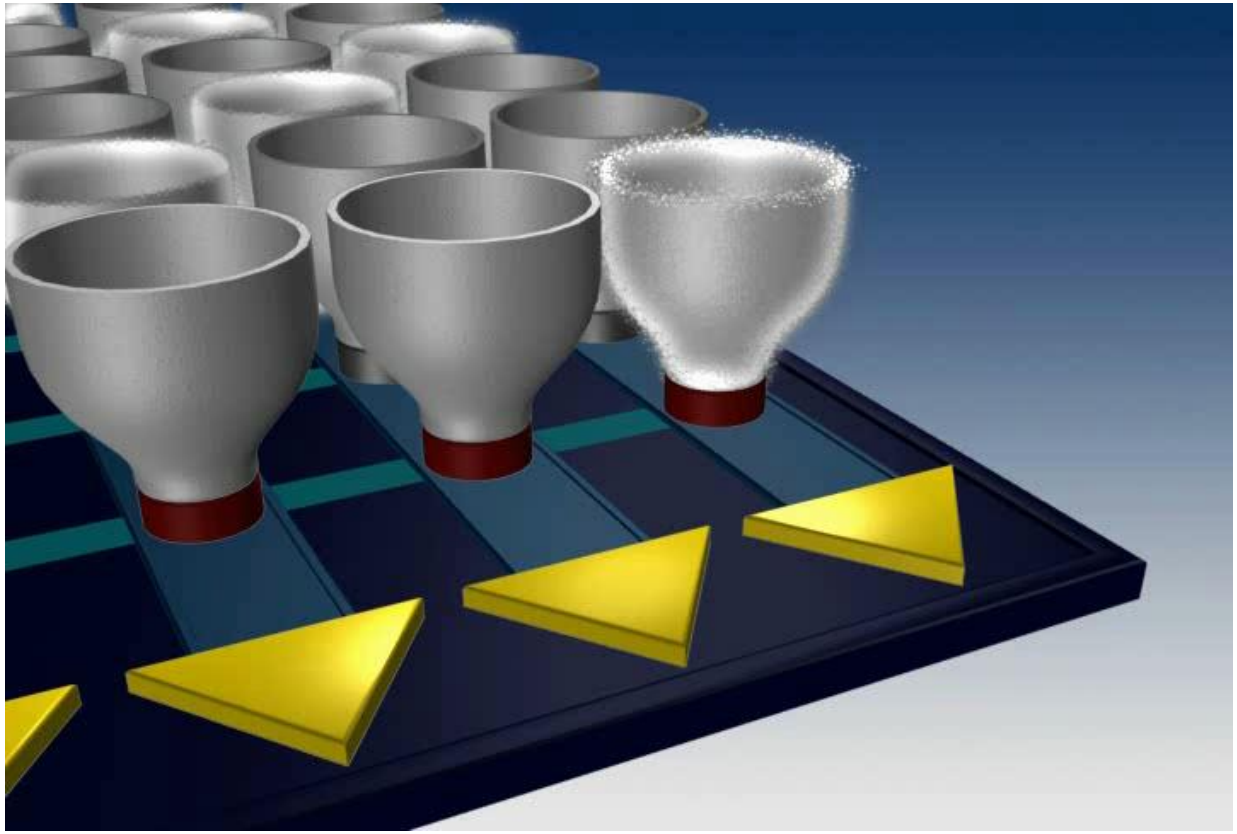
- In 1966, Dr. Robert Dennard at IBM Thomas J Watson Research Center invented the “Field Effect Transistor Memory” – now called DRAM
- One transistor per bit







DRAM Operation



More or Less of Moore

- Is Moore's Law obsolete?
  - Economic realities
  - New platforms drive technology
- DRAM cell physics
  - Cell and bit line capacitance
  - Access device performance
- Economics
  - Masks
  - Lithography
  - Process complexity

DRAM Process Complexity

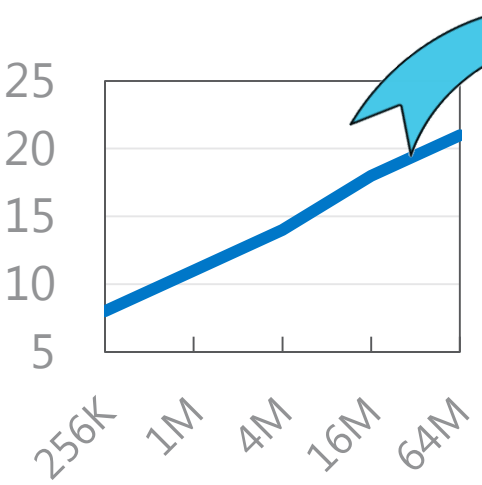


Figure 1
1999 Data showing DRAM mask count trends.

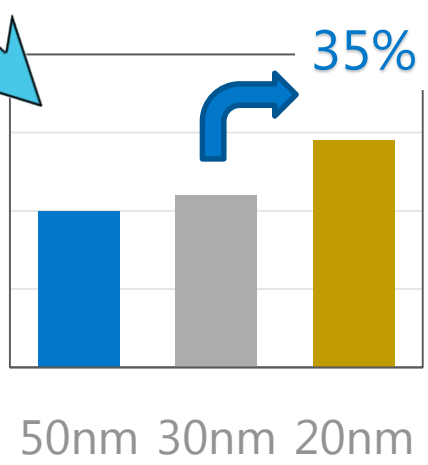


Figure 2
2015 Data showing DRAM mask levels with process nodes.

DRAM Process Complexity

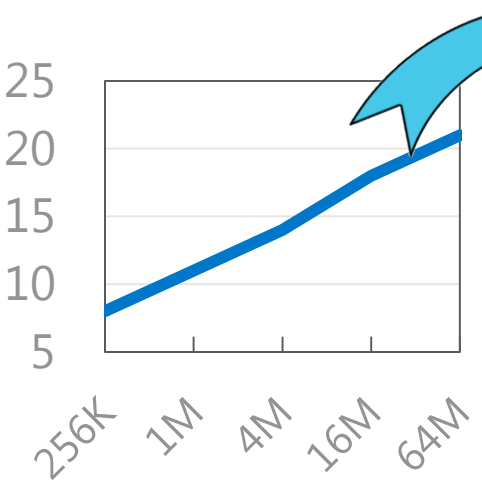


Figure 1
1999 Data showing DRAM mask count trends.

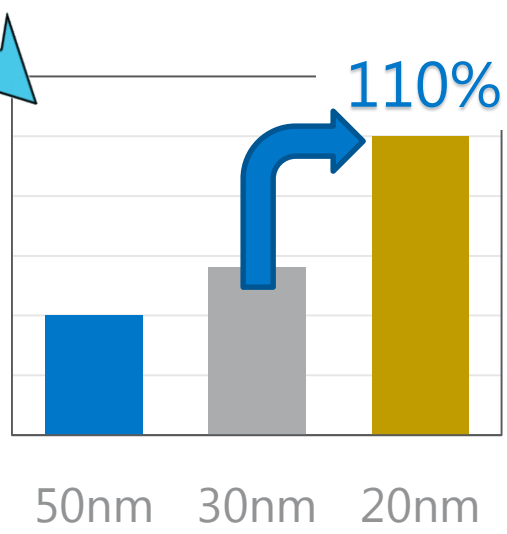


Figure 3
2015 Data showing DRAM non-litho steps per critical mask level with process nodes.

DRAM Process Complexity

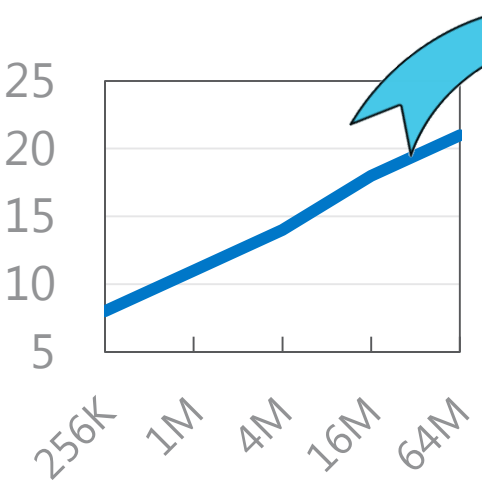


Figure 1
1999 Data showing DRAM mask count trends.

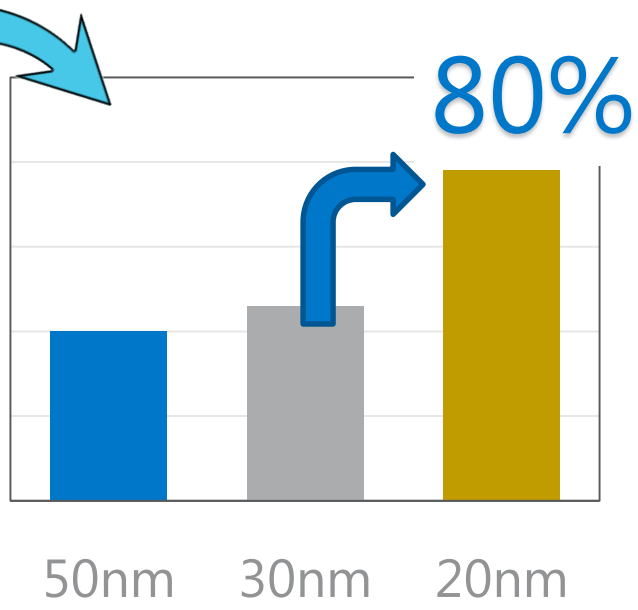
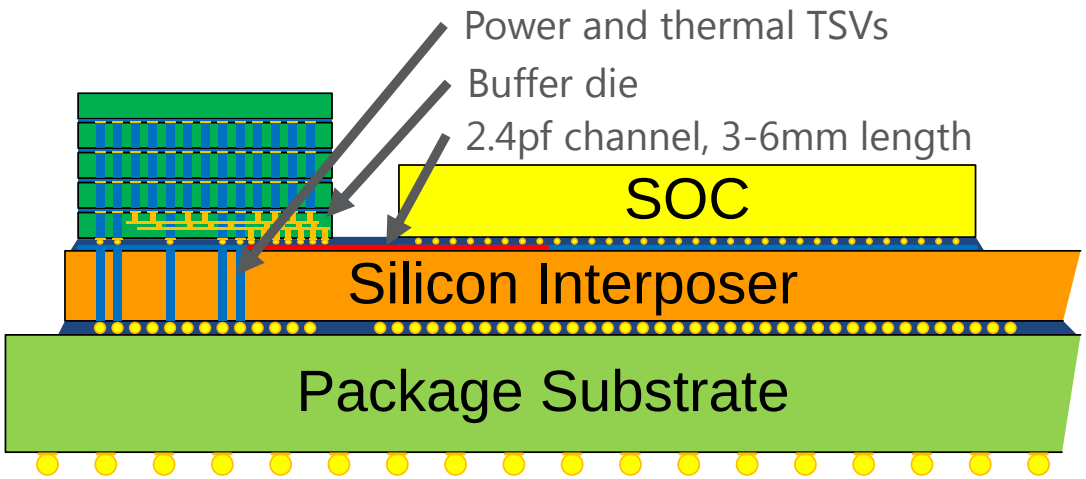


Figure 4
2015 Data showing DRAM cleanroom space per wafer out with process nodes.

Solution: System Innovation

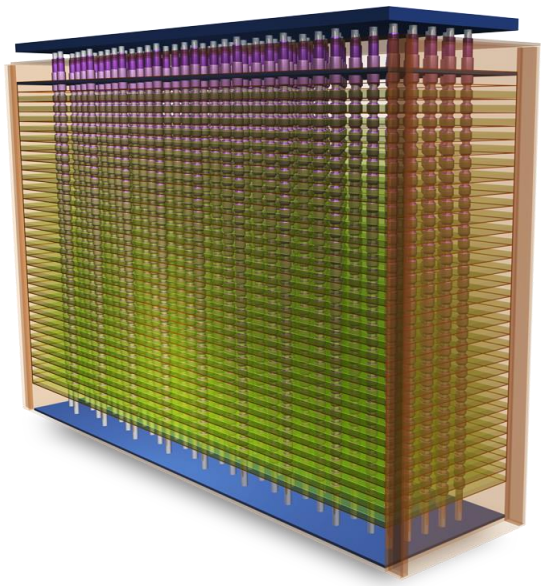
3D DRAM Solutions: HBM & HMC

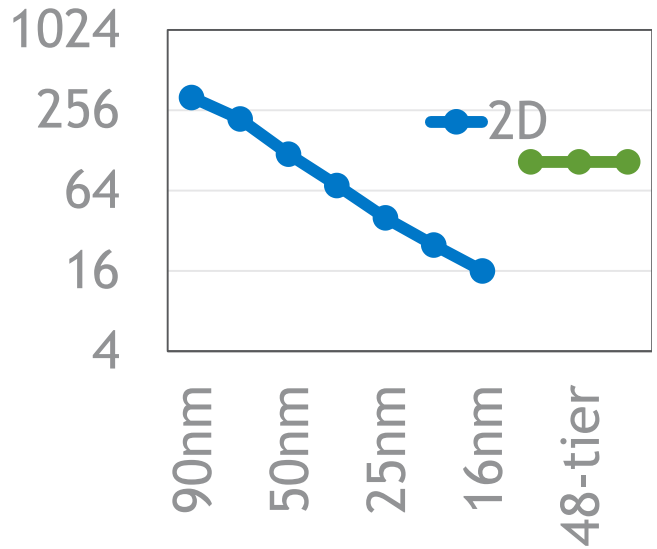


HBM Micro Pillar Area	
Area	2.88 x 6.05mm
Bump count	6610
I/O	1632
VSS	1048
VDD	480
VDDQ	432
VPP	160
DA (test)	60

3D NAND

- Relaxed pitch gives better SNR
- Allows for “More than Moore” growth in density





Electrons per cell versus node

Optimizing for Applications

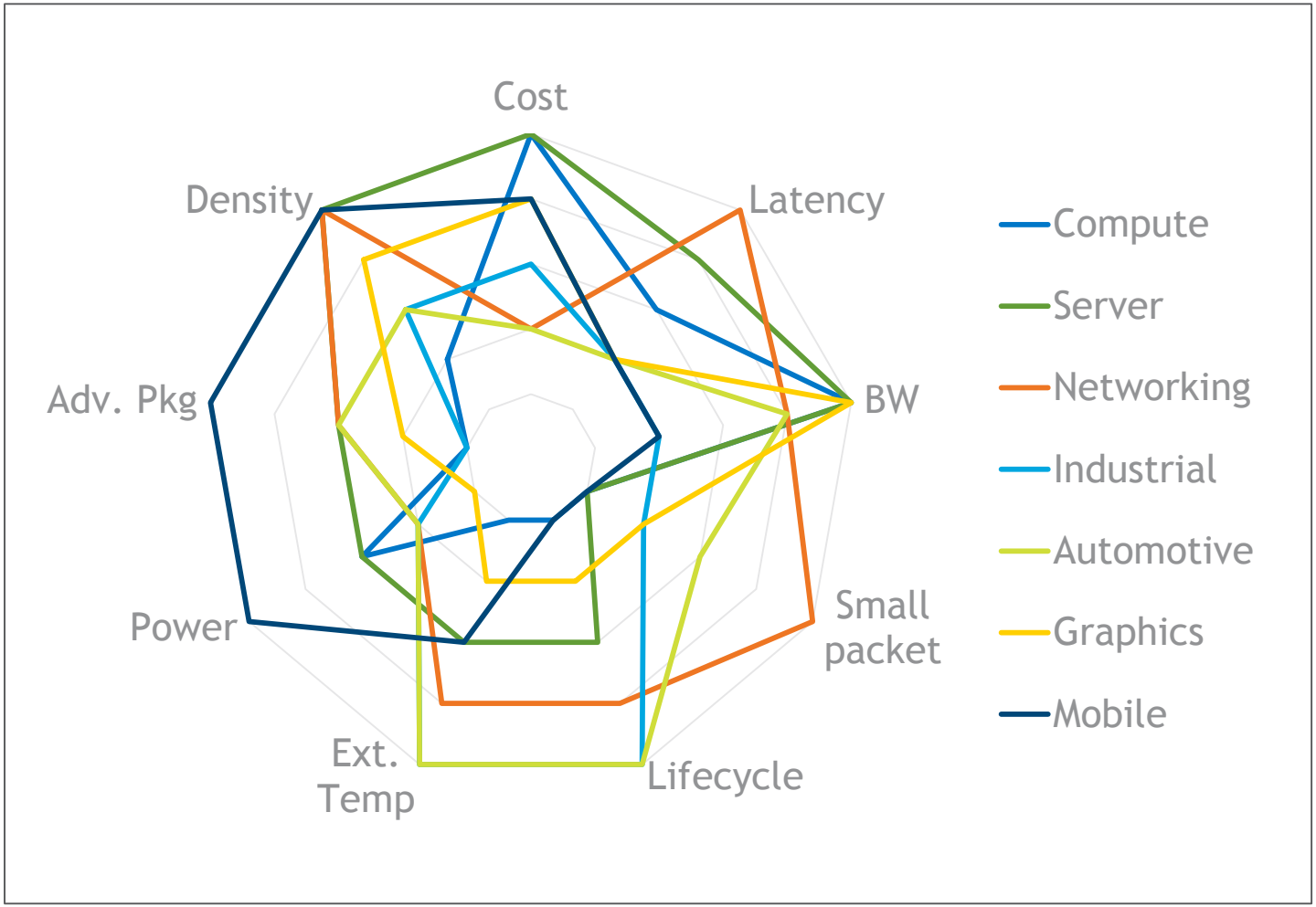
- The personal computer drove DRAM directions.
- Smart phones and tablets have driven NAND directions.
- In memory, “One size fits all” does not apply.

Applications:
Compute
Server
Networking
Industrial
Automotive
Graphics
Mobile
New?

Drivers

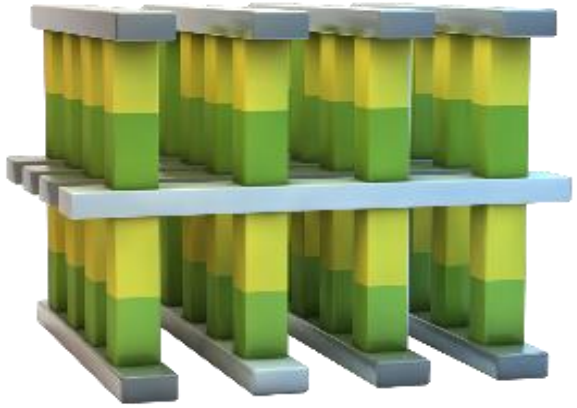
- Compute: “Tablification”
- Server: Integration of storage, 44 Zettabytes of data
- Networking: 60% more traffic/year, 30% less energy
- Industrial: IoT, M2M, Automation
- Automotive: Autonomous/assistive vehicles
- Graphics: Accelerate this…
- Mobile: 28 Exabytes of mobile traffic in 2015

Application “Care-Abouts”



Emerging Memories

- 3D Xpoint™
- MRAM
- Other?



3D Xpoint™ 2-layer memory array

Criteria:

- Cost
- Endurance
- Non-Volatility
- Manufacturability
- Is it an addition to the hierarchy or a replacement?

Opportunities

- Innovative architectures and interfaces
- Innovative packaging, 3D integration
- Optimization for applications, i.e. low latency
- Embrace emerging technology
- Teach an old dog new tricks

Questions?


Micron®