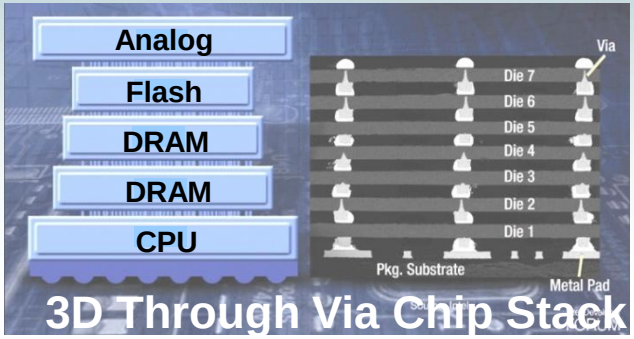
3D IC Overview

Robert Patti, CTO

rpatti@tezzaron.com

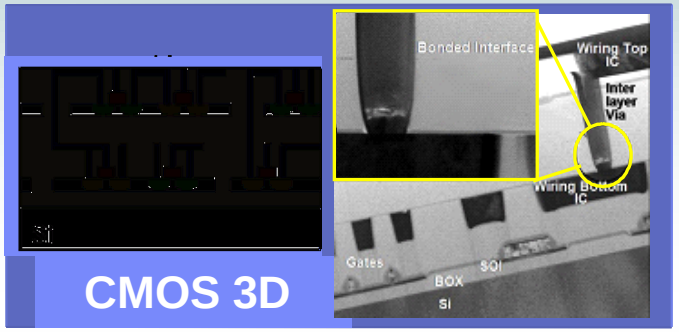
Span of 3D Integration

Packaging



IBM/Samsung

Wafer Fab



IBM

Tezzaron 3D-ICs

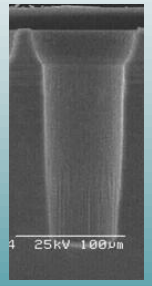
100-1,000,000/sqmm

1000-10M Interconnects/device

1s/sqmm

Peripheral I/O

- Flash, DRAM
- CMOS Sensors



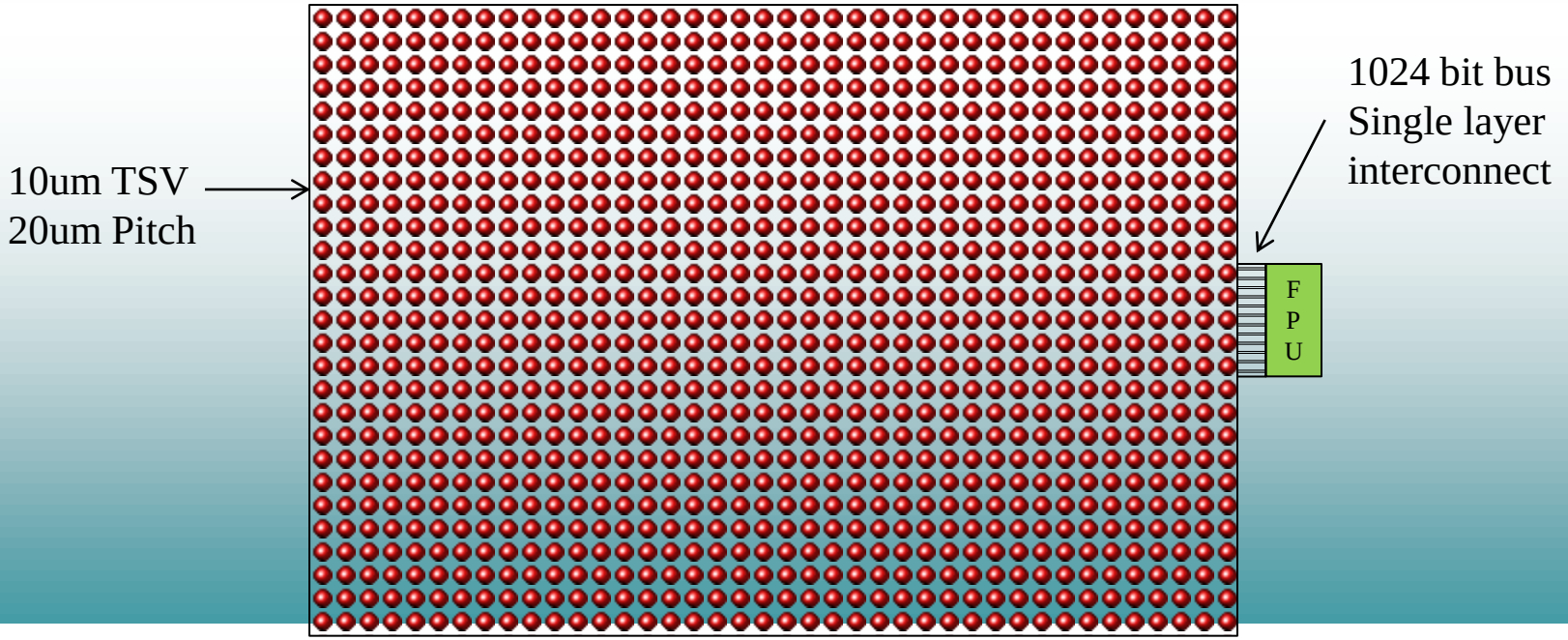
100,000,000s/sqmm

Transistor to Transistor

- Ultimate goal

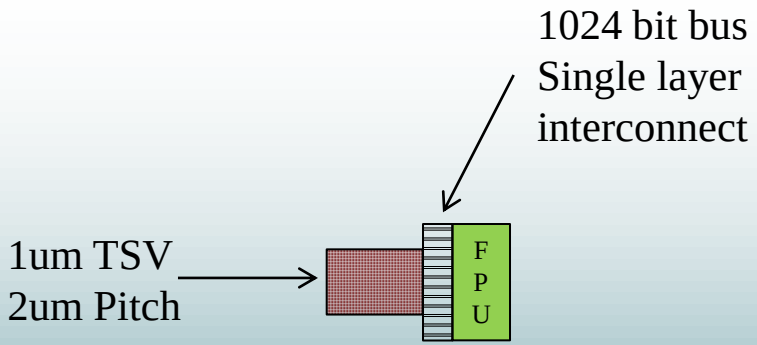
TSV Pitch ≠ Area ÷ Number of TSVs

- TSV pitch issue example
  - 1024 bit busses require a lot of space with larger TSVs
  - They connect to the heart and most dense area of processing elements
  - The 45nm bus pitch is ~100nm; TSV pitch is >100x greater
  - The big TSV pitch means TOF errors and at least 3 repeater stages



TSV Pitch ≠ Area ÷ Number of TSVs

- TSV pitch issue example
  - 1024 bit busses require a lot of space with larger TSVs
  - They connect to the heart and most dense area of processing elements
  - The 45nm bus pitch is ~100nm; TSV pitch is >100x greater
  - The big TSV pitch means TOF errors and at least 3 repeater stages



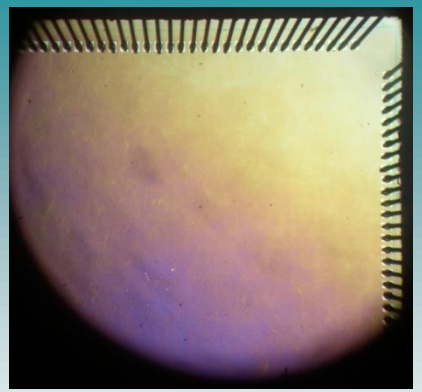
3D Interconnect Characteristics

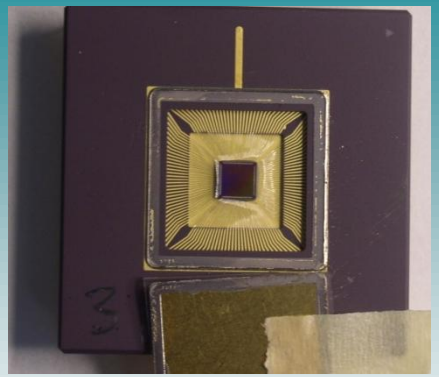
	SuperContact™ I 200mm Via First, FEOL	SuperContact™ III 200mm Via First, FEOL	SuperContact™ IV 200mm Via First, FEOL	Interposer TSV	Bond Points	Die to Wafer
Size L X W X D Material	1.2 µ X 1.2 µ X 6.0µ W in Bulk	0.85 µ X 0.85 µ X 10µ W in Bulk	0.60 µ X 0.60 µ X 2µ W in SOI	10 µ X 10 µ X 100 µ Cu	1.7 µ X 1.7 µ Cu	3 µ X 3 µ Cu
Minimum Pitch	<2.5 µ	1.75 µ	1.2 µ	30/120 µ	2.4 µ	5 µ
Feedthrough Capacitance	2-3fF	3fF	0.2fF	250fF	<<	<25fF
Series Resistance	<1.5 Ω	<3 Ω	<1.75 Ω	<0.5 Ω	<	<

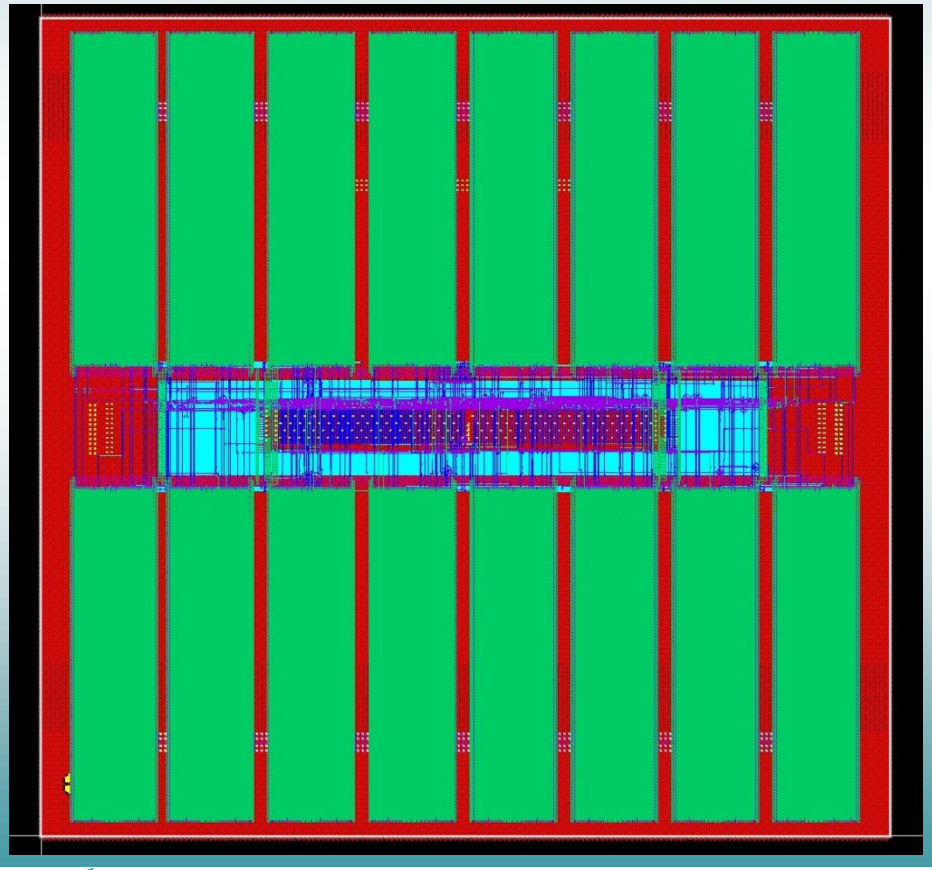
Small fine grain TSVs are fundamental to 3D enablement

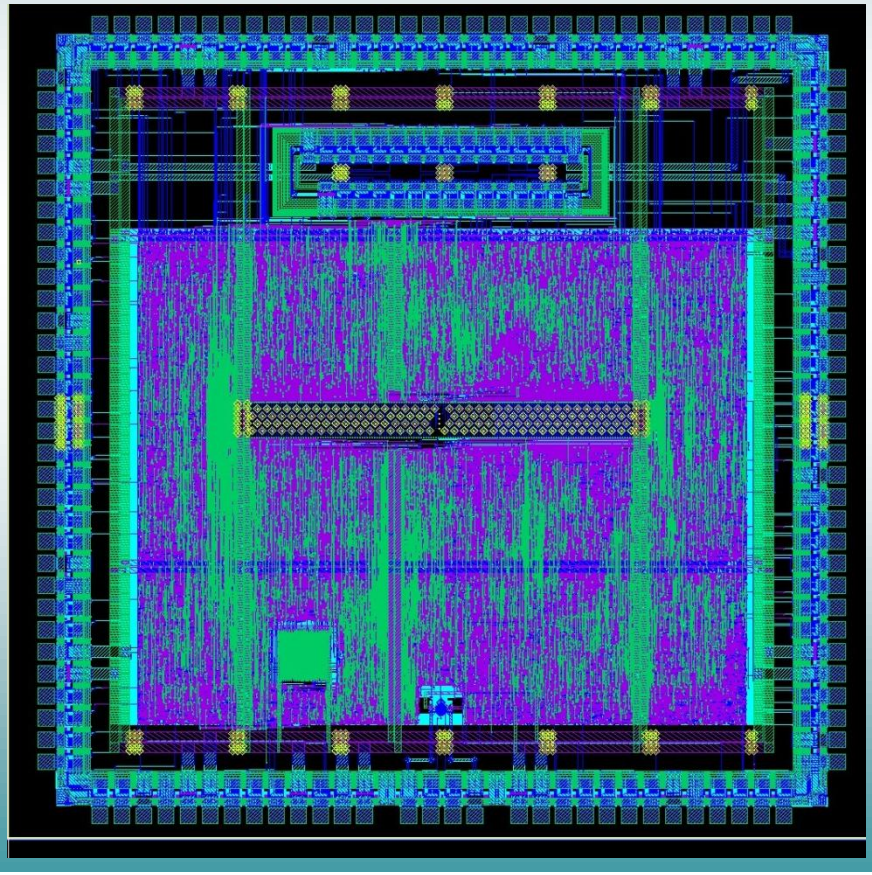
R8051/Memory

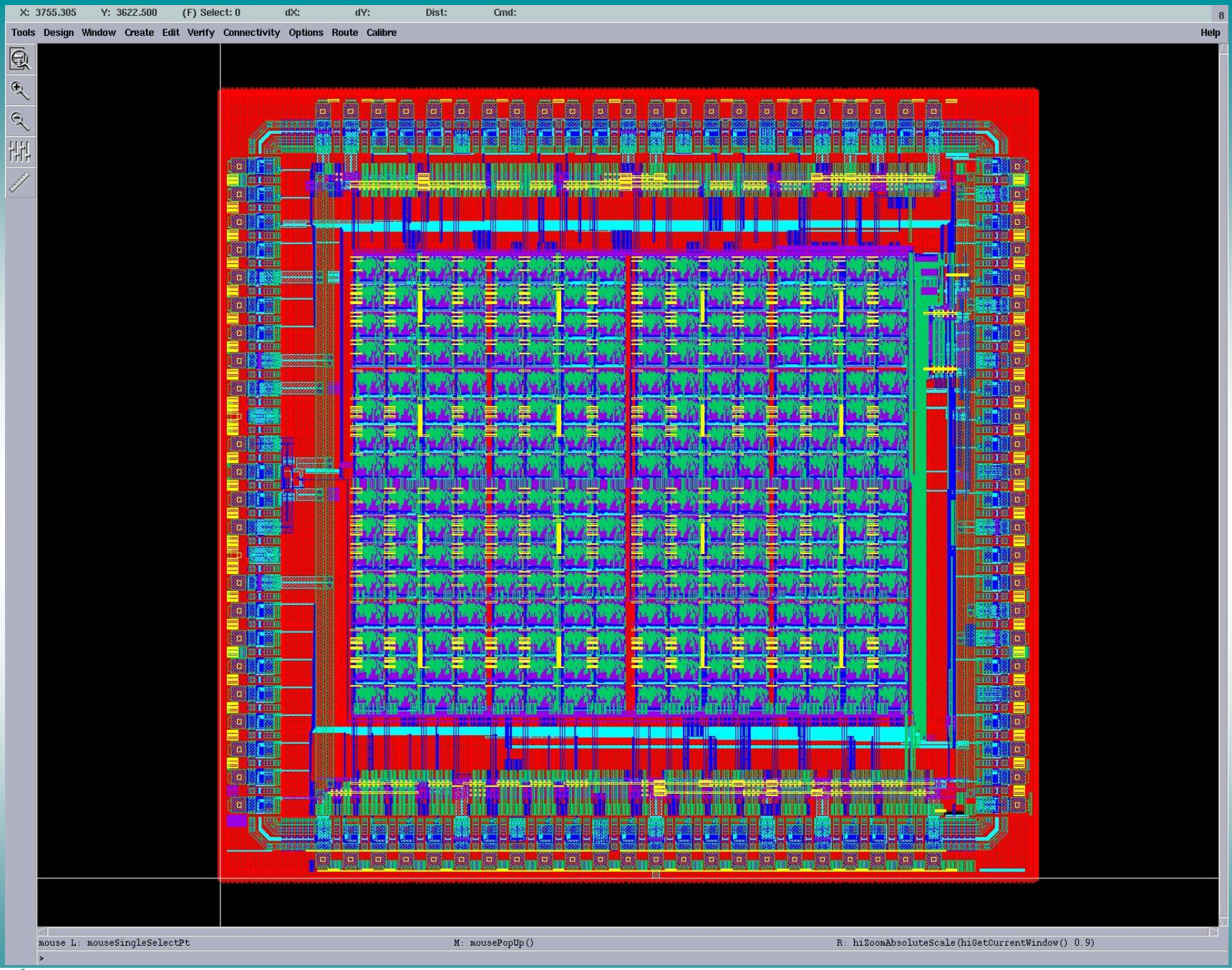
5X Performance
1/10th Power



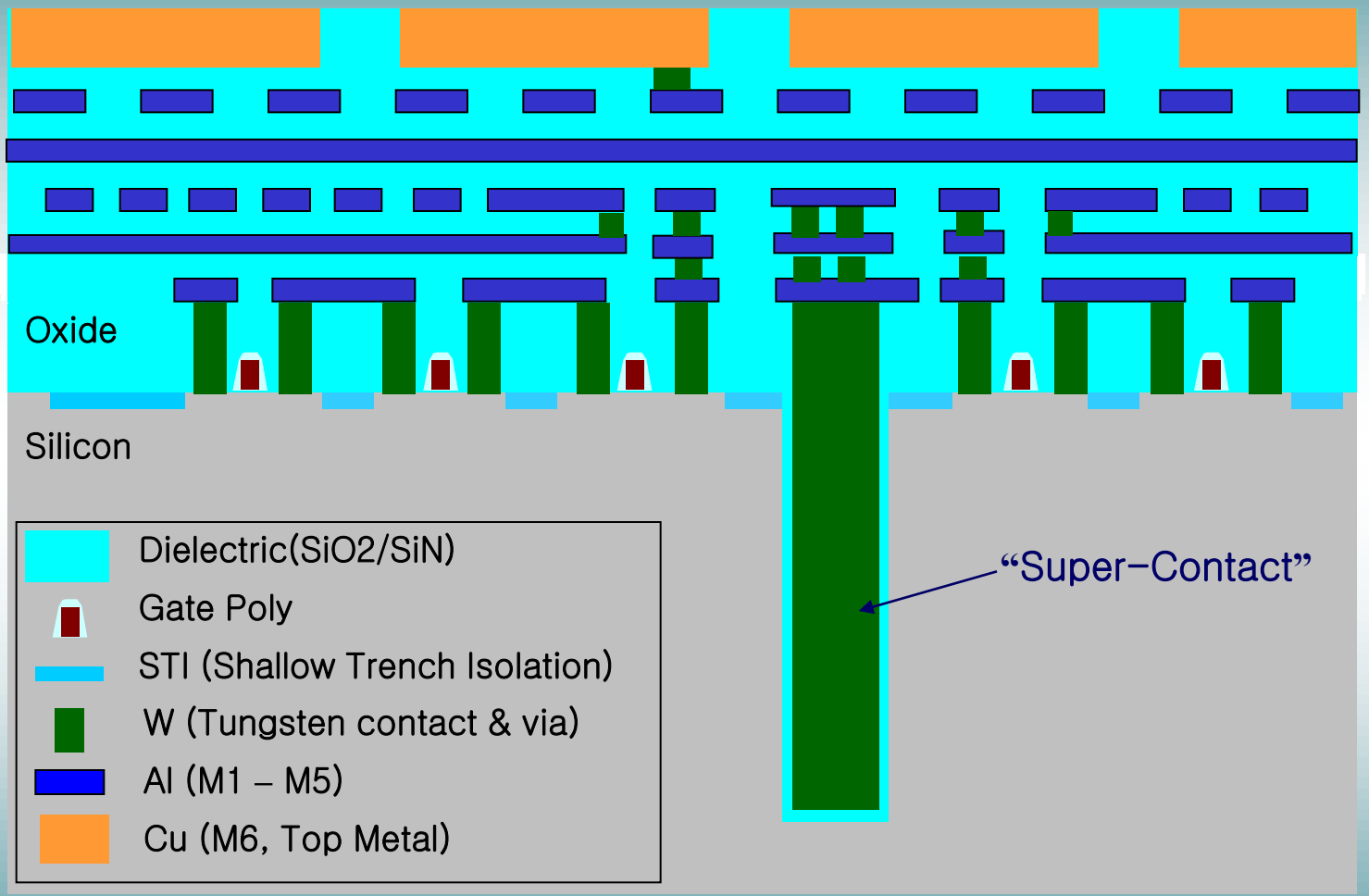




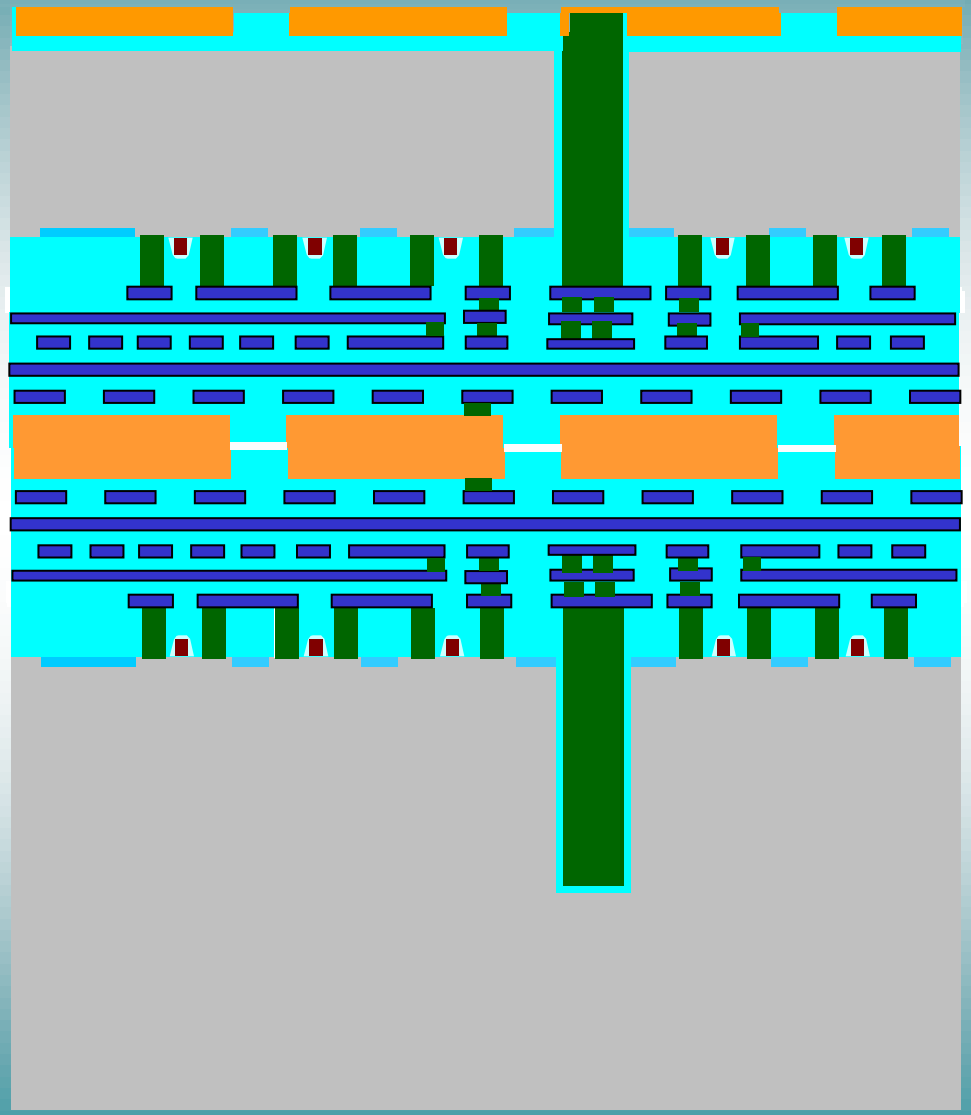



X: 3755.305 Y: 3622.500 (F) Select: 0 dX: dY: Dist: Cmd:
Tools Design Window Create Edit Verify Connectivity Options Route Calibre Help
mouse L: mouseSingleSelectPt M: mousePopUp() R: hiZoomAbsoluteScale(hiGetCurrentWindow() 0.9)

A Closer Look at Wafer-Level Stacking



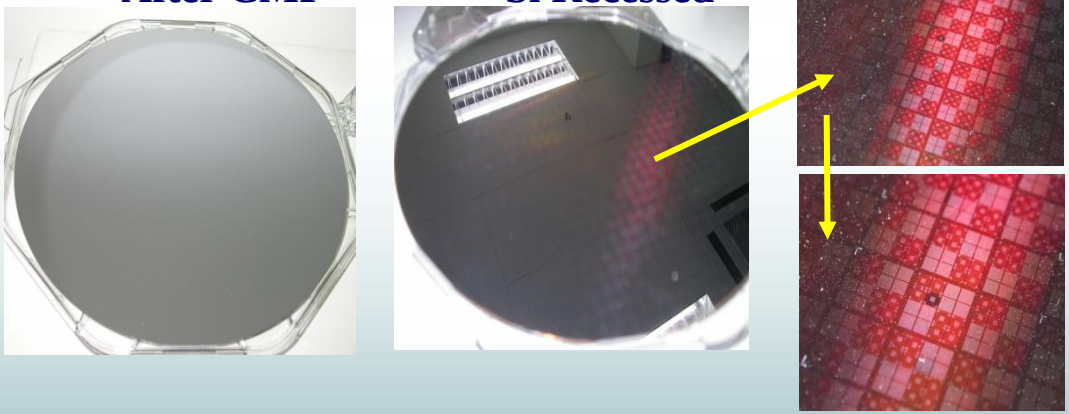
Next, Stack a Second Wafer & Thin:



Stacking Process Sequential Picture

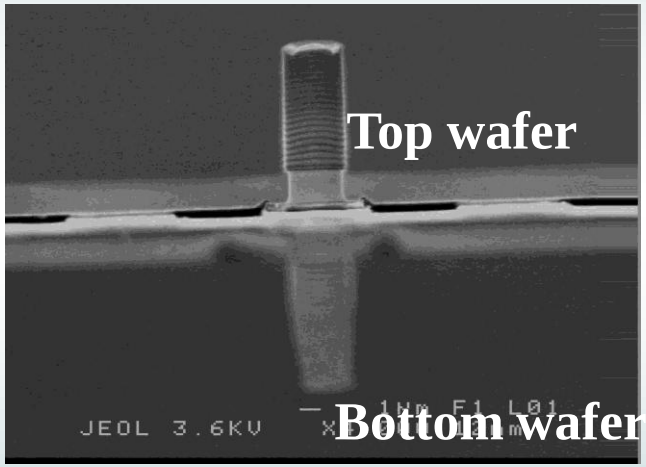
Two wafer Align & Bond → Course Grinded → Fine Grinded

→ After CMP → Si Recessed

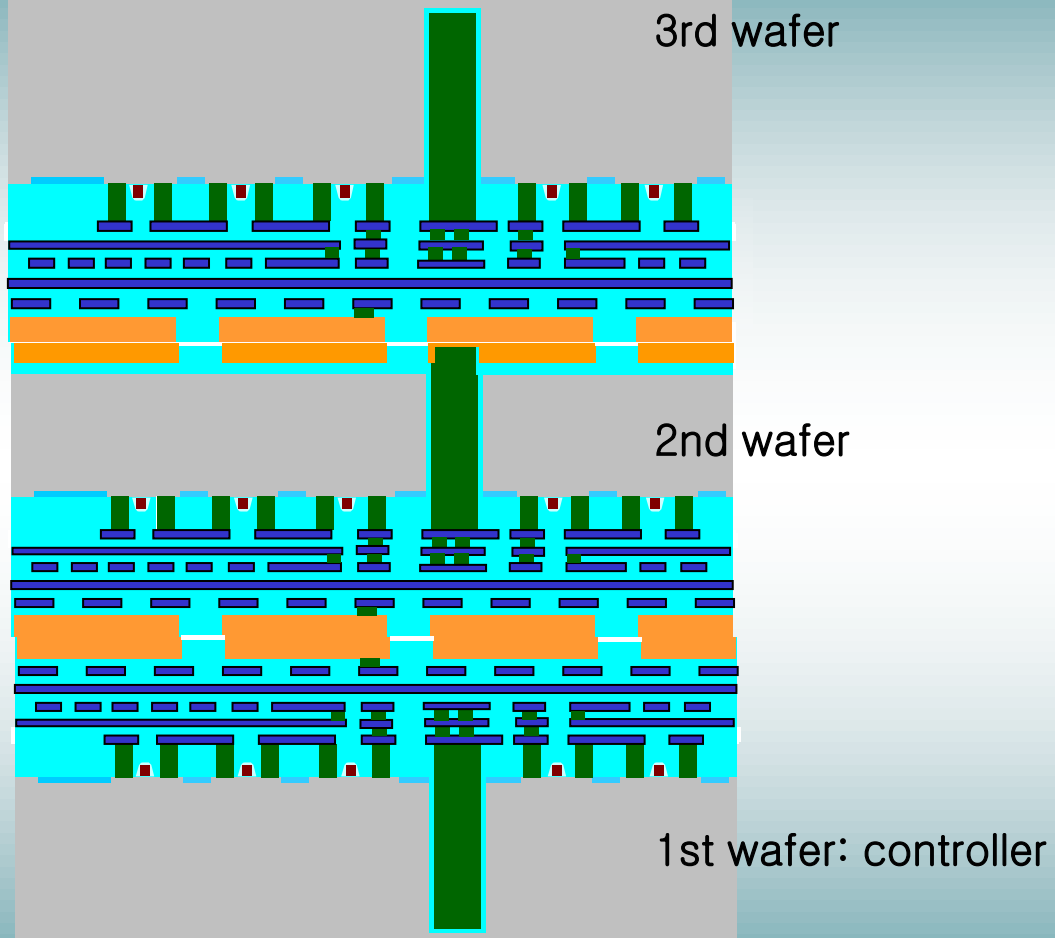


High Precision Alignment

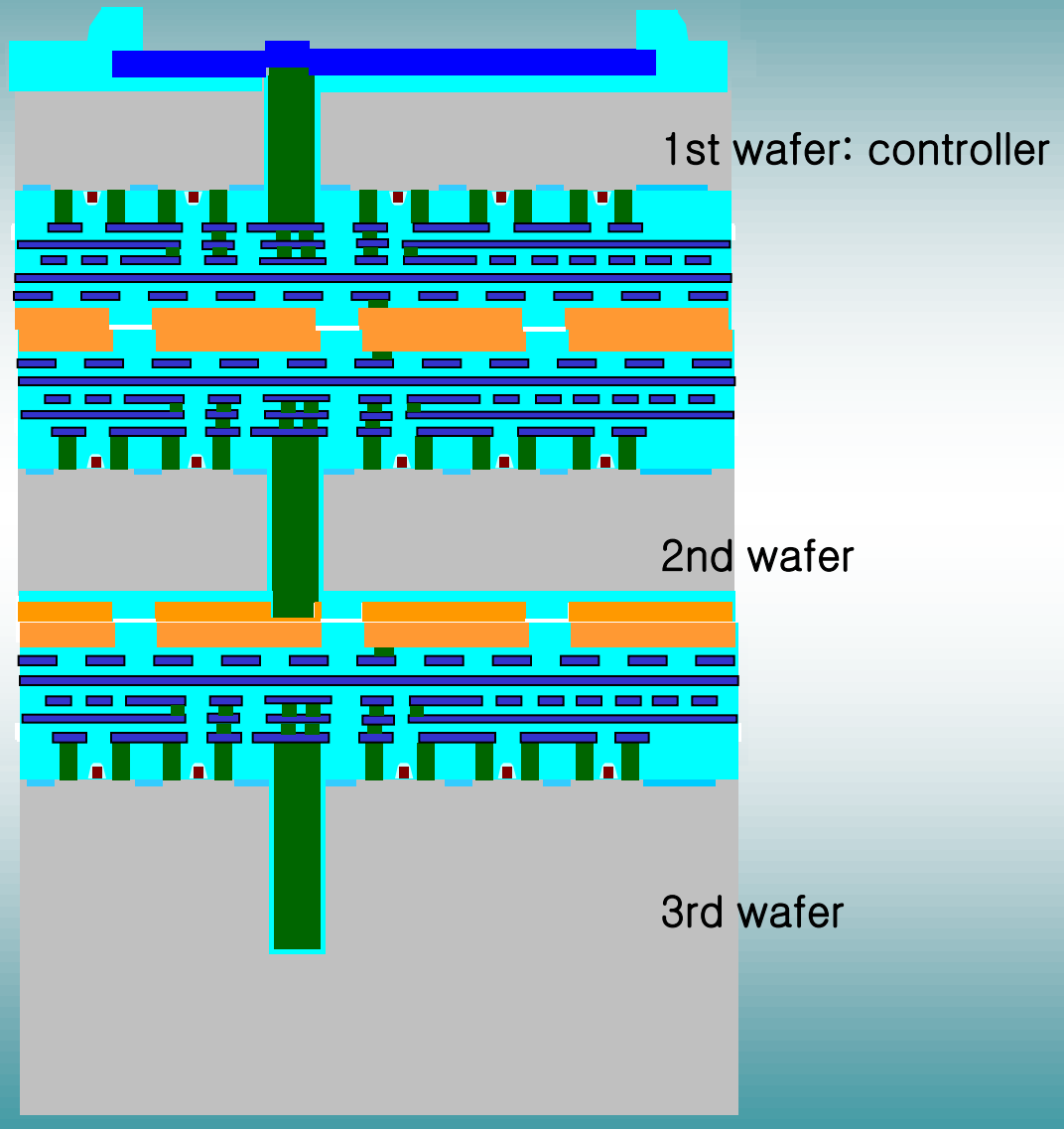
Misalign=0.3um



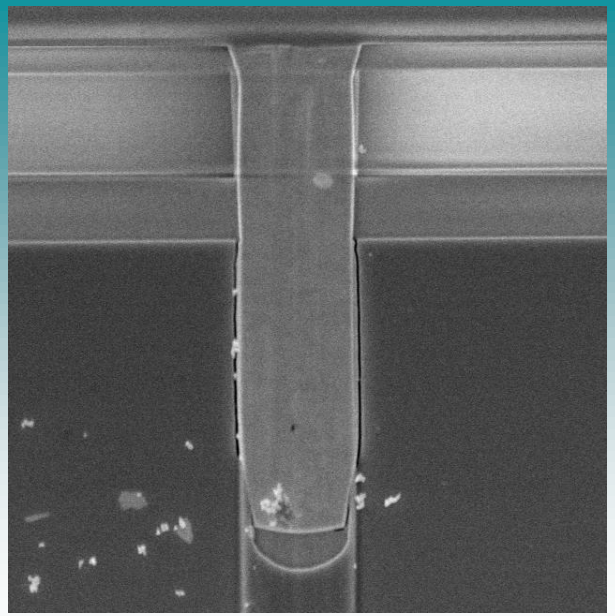
Then, Stack a Third Wafer:

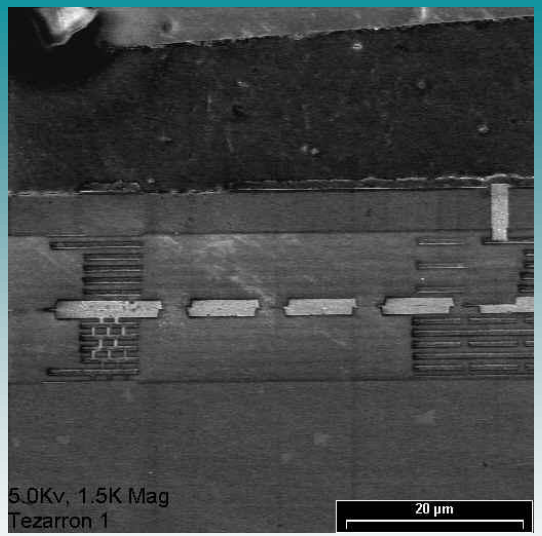


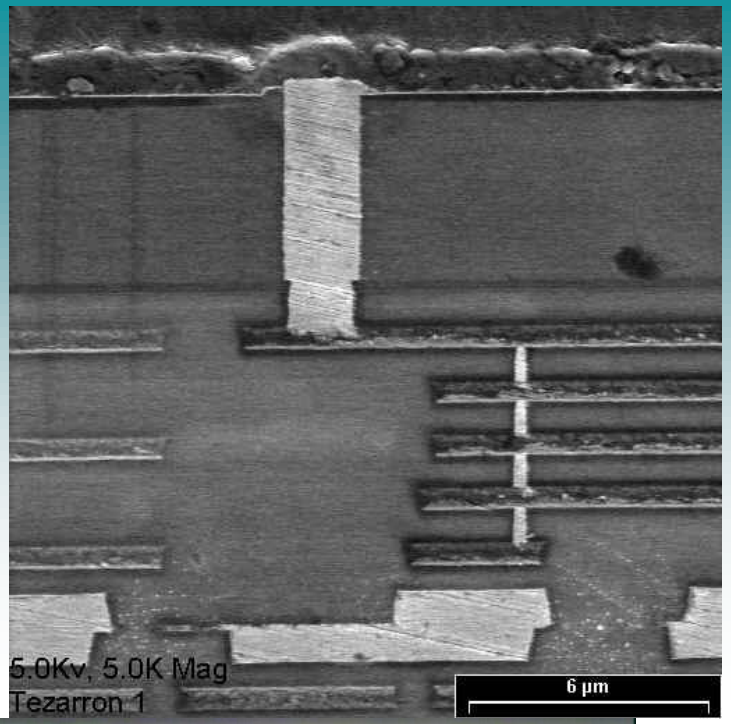
Finally, Flip, Thin & Pad Out:

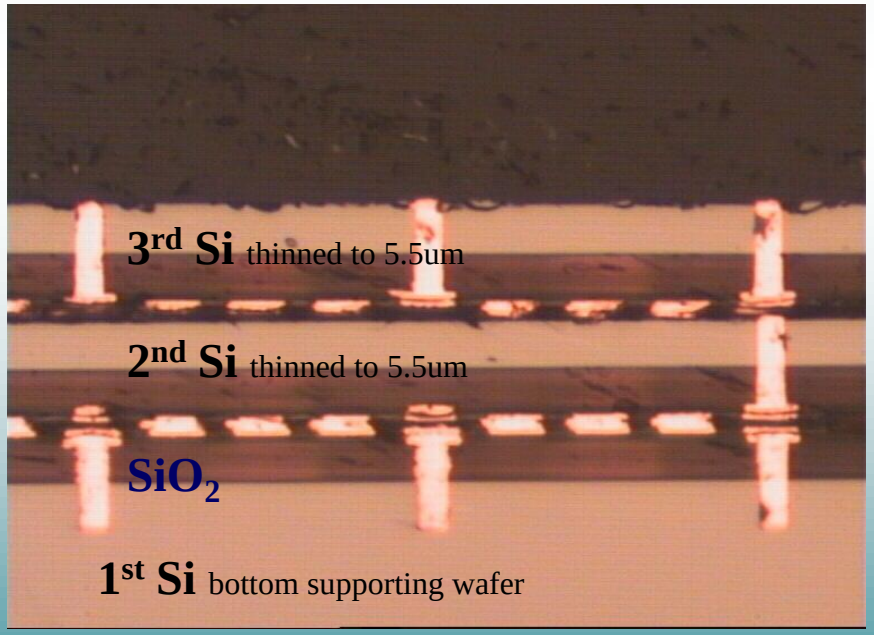


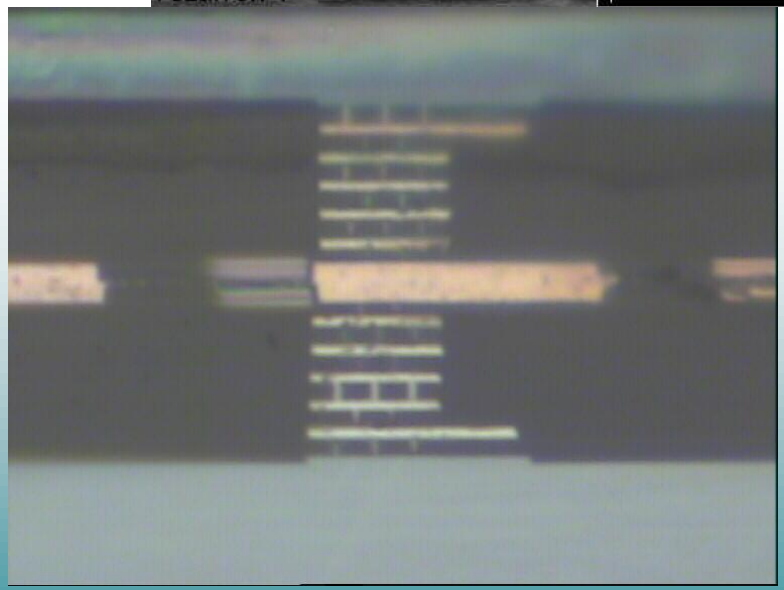
This is the completed stack!

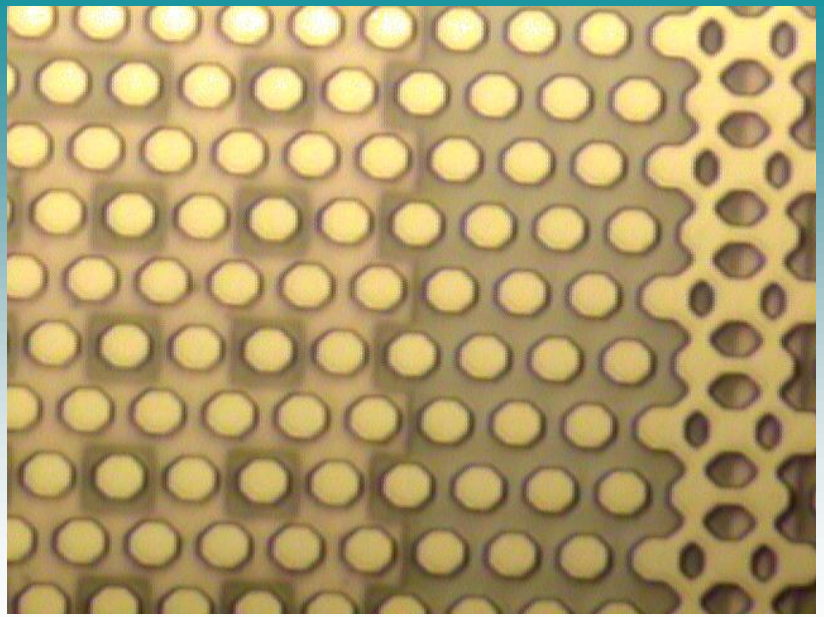


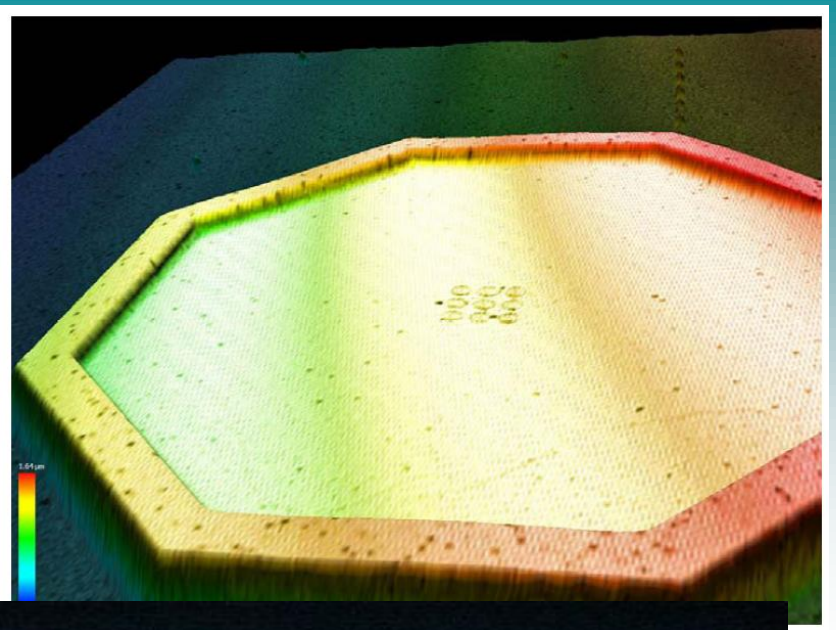

5.0Kv, 1.5K Mag
Tezarron 1
20 µm

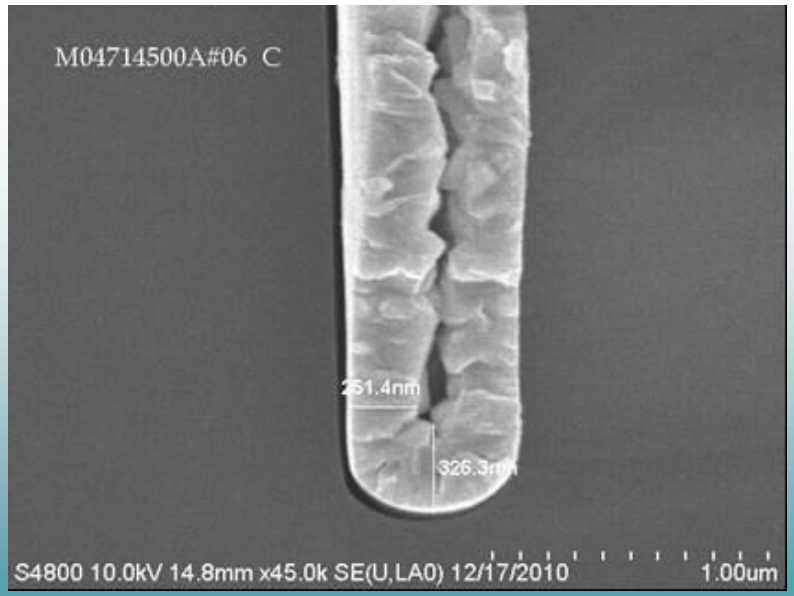

5.0Kv, 5.0K Mag
Tezarron 1
6 µm

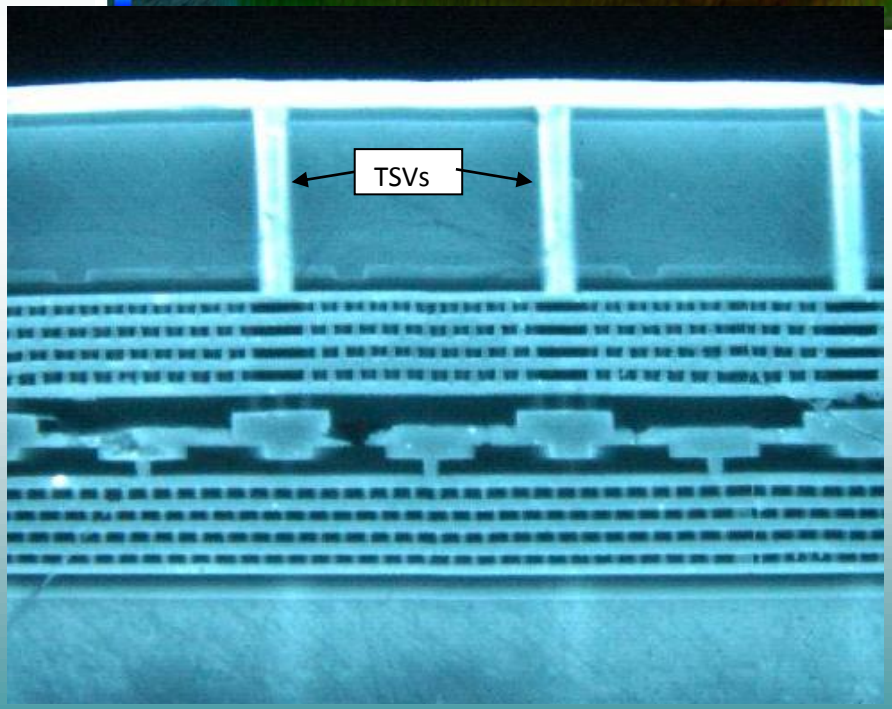

3rd Si thinned to 5.5um
2nd Si thinned to 5.5um
SiO2
1st Si bottom supporting wafer



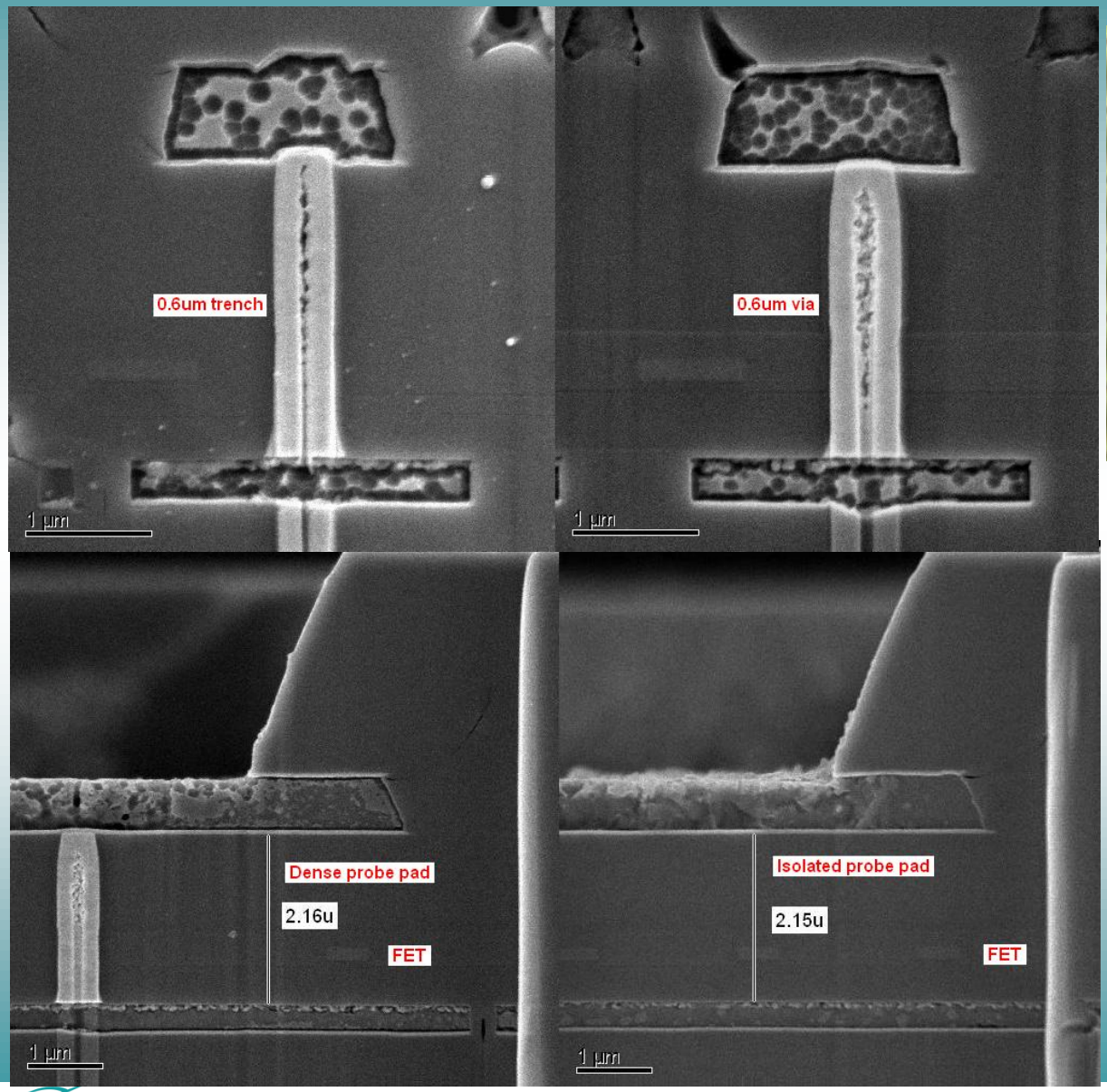


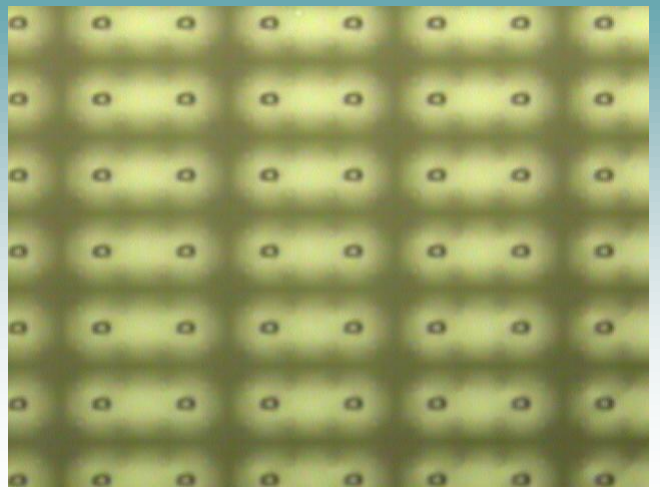



M04714500A#06 C
251.4nm
326.3nm
S4800 10.0kV 14.8mm x45.0k SE(U,LA0) 12/17/2010
1.00um

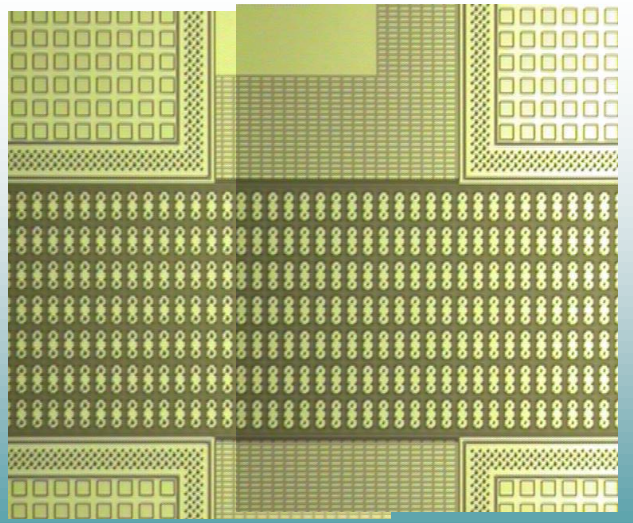

TSVs

0.6um SOI TSV





120K TSVs



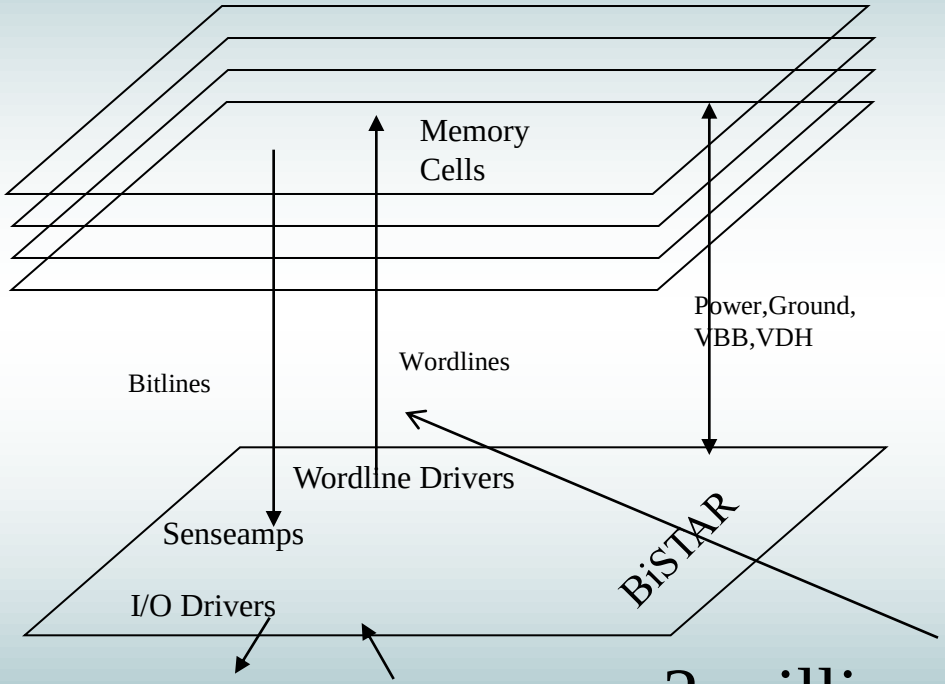
DRAM wants 2 different processes!

Bit cells	Low leakage -slow refresh -low power -low GIDL	High Vt Devices Vneg Well Thick Oxide
Sense Amps Word line drivers Device I/O	High speed -better sensitivity -better bandwidth -lower voltage	Low Vt Devices Copper interconnect Thin Oxides

“Dis-Integrated” 3D Memory

Memory Layers from DRAM fab

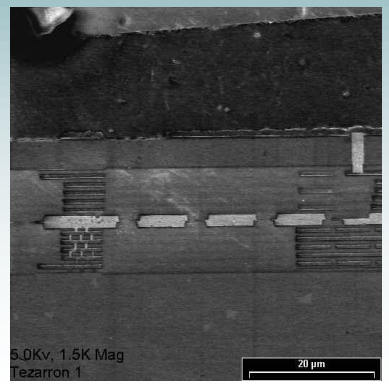
Controller Layer from high speed logic fab



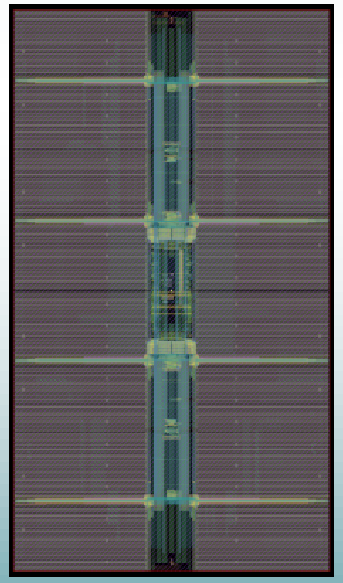
2 million vertical connections per lay per die

Better yielding than 2D equivalent!

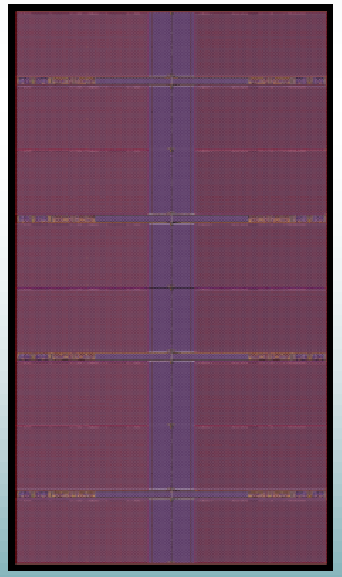
Octopus Stack



2 Layer Stacked Device
(SEM)



DRAM Control/Logic



DRAM Memory Cells

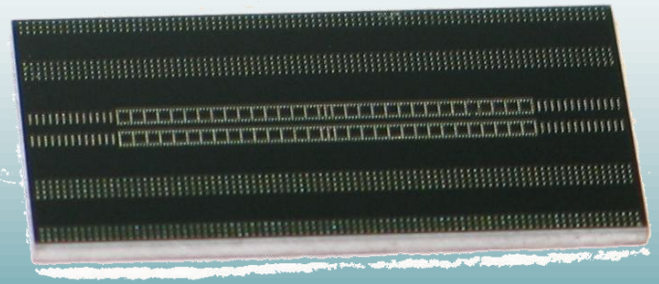
3D DRAMs

Octopus I

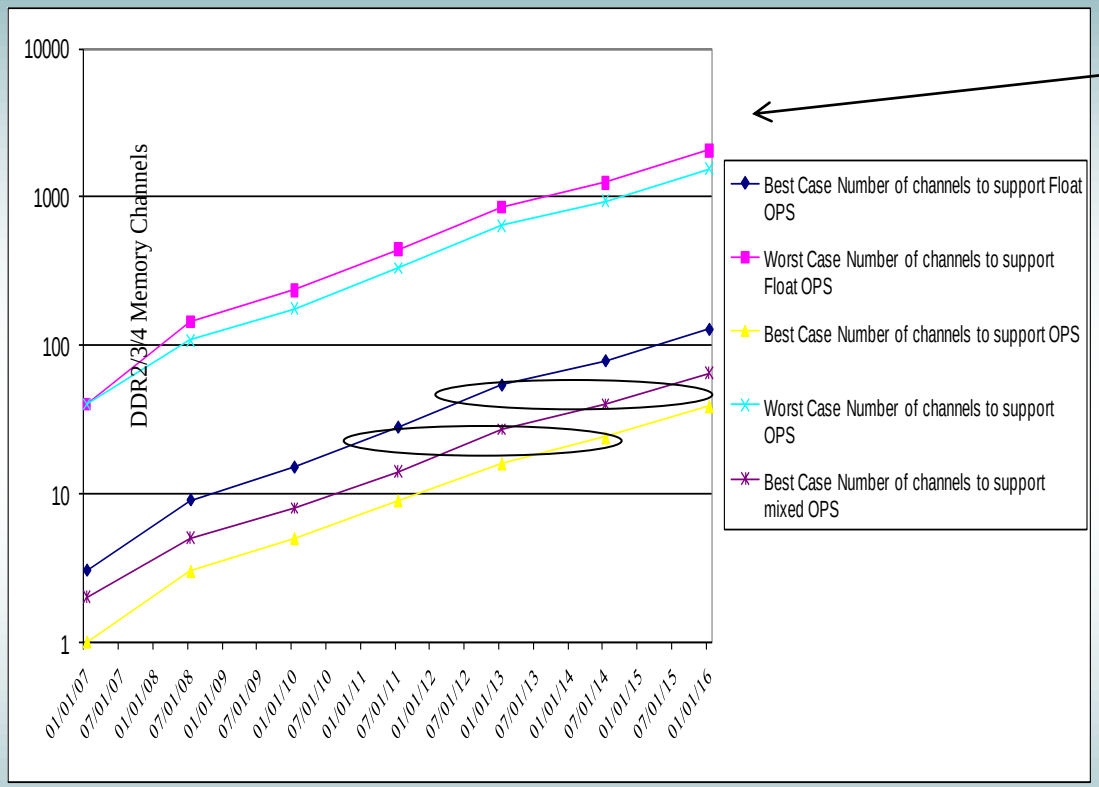
- 1-4Gb
- 16 Ports x 128bits (each way)
- @1GHz
  - CWL=0 CRL=2 SDR format
  - 5ns closed page access to first data (aligned)
  - 12ns full cycle memory time
  - 288GB/s data transfer rate
- Max clk=1.6GHz
- Internally ECC protected, Dynamic self-repair, Post attach repair
- 115C die full function operating temperature
- JTAG/Mailbox test&configuration

Octopus II

- 4-64Gb
- 64-256 Ports x 64bits (each way)
- @1GHz
  - 5-7ns closed page access to first data (aligned)
  - 12ns full cycle memory time
  - 2TB/s data transfer rate



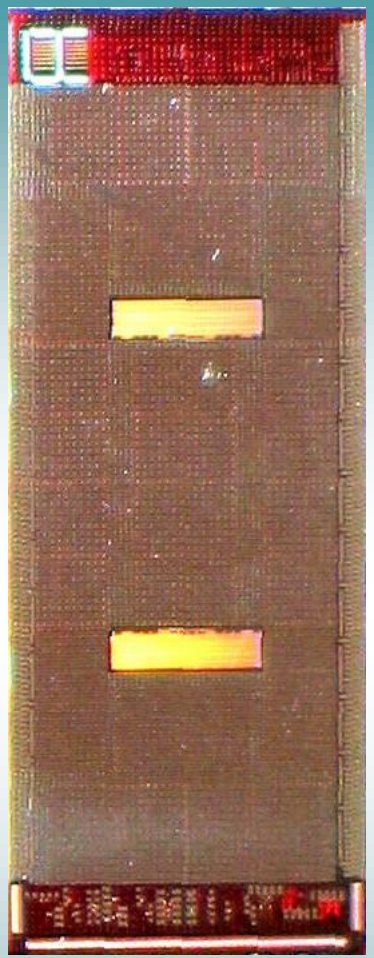
The Industry Issue



Need 50x bandwidth improvement.
Need 10x better cost model than embedded memory.

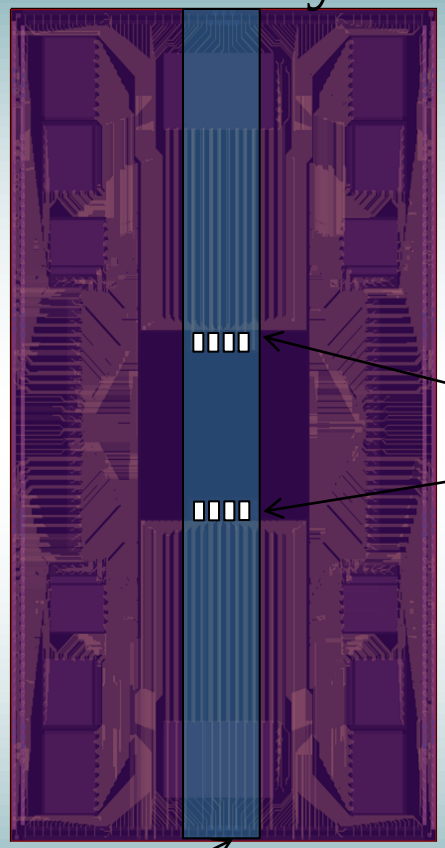
- To continue to increase CPU performance, exponential bandwidth growth required.
- More than 200 CPU cycles of delay to memory results in cycle for cycle CPU stalls.
- 16 to 64 Mbytes per thread required to hide CPU memory system accesses.
- No current extension of existing IC technology can address requirements.
- Memory I/O power is running away.

Logic on Memory

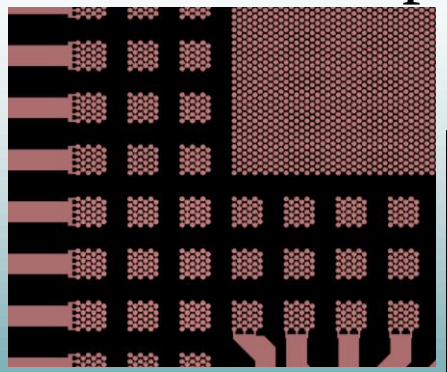


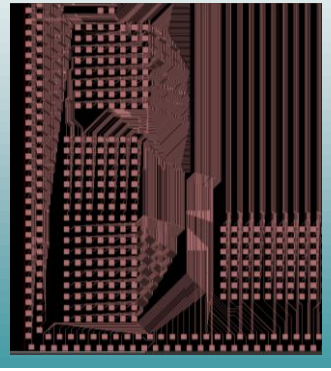
Logic

Memory



8 DRAM ports
16x21 pad array





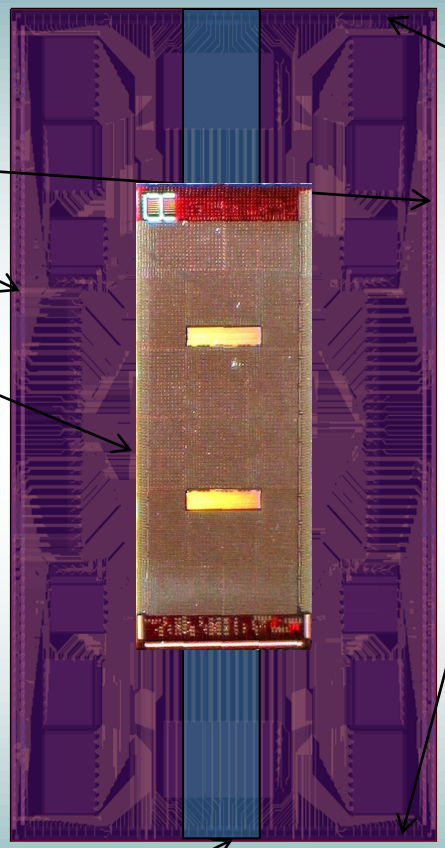
>10µf bypass caps
SS ~4,000pf

Logic on Memory

172 pads

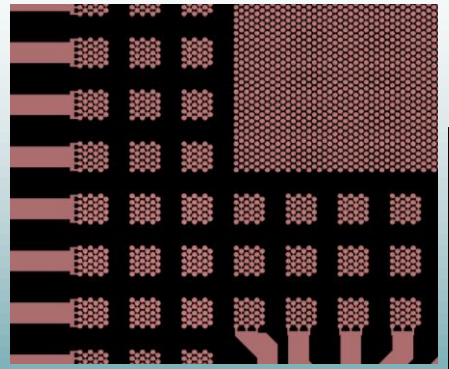
199 I/O Bondpoints/side

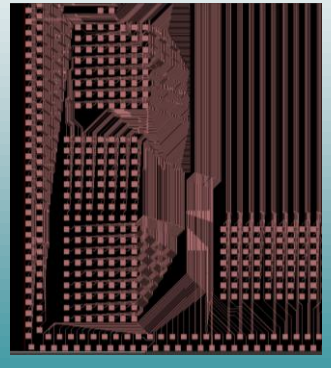
Memory also acts as interposer



92 pads

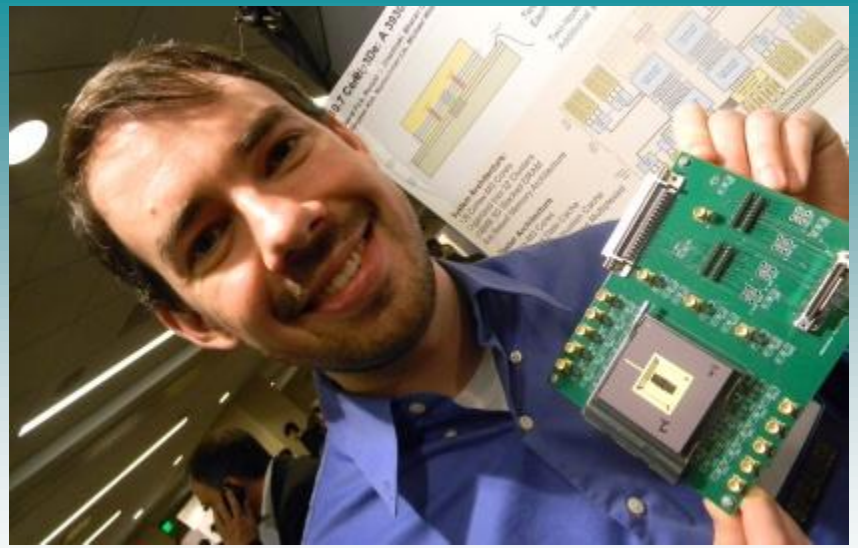
(528 total pads at edge, stagger
250um pad, 125um pitch
~1500 available pads)





>10µf bypass caps

SS ~4,000pf



David Fick of the University of Michigan showed Centip3De, a 3-D IC stack using 128 ARM Cortex M3 cores and 256 Mbytes of stacked DRAM operating at near threshold voltage. (Paper 10.7)

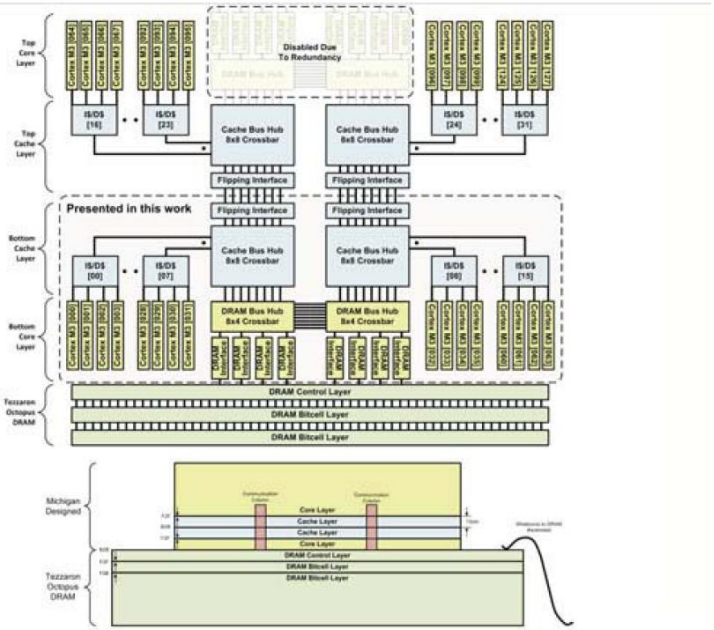
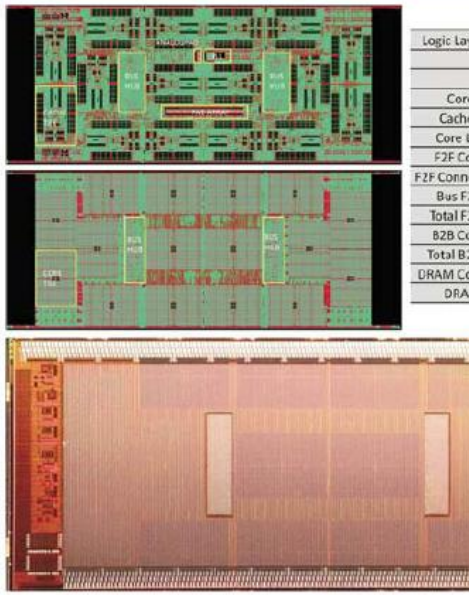


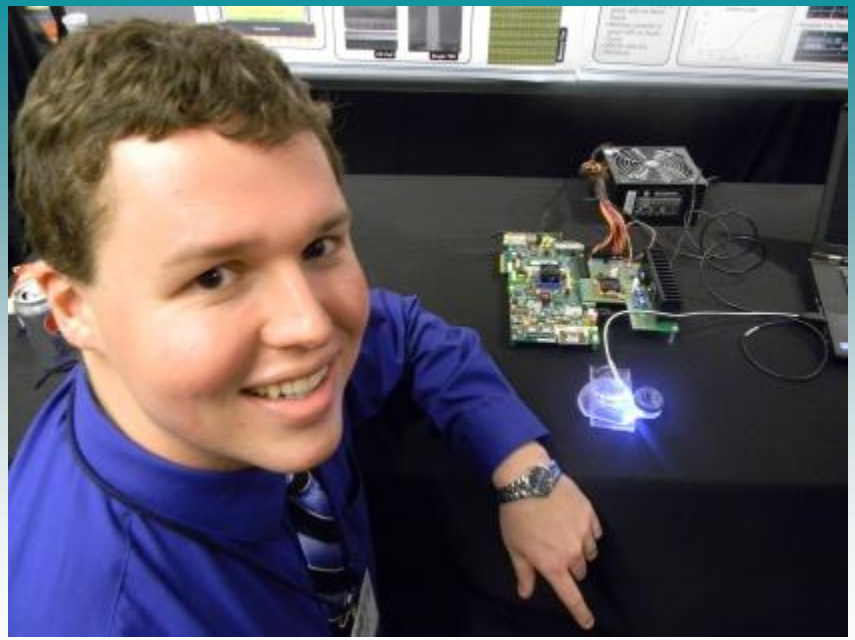
Figure 10.7.3: System block diagram and system side view. The fabricated system has an unthinned cache layer and a thinned core layer, with wirebonds connecting to TSVs on the backside.



Logic Layer Dimensions	12.66 x 5mm
Technology	130nm
Metal Layers	5
Core Layer Devices	28.4 M
Cache Layer Devices	18.0 M
Core Layer Thickness	12um
F2F Connection Pitch	5 um
F2F Connections/Cluster	1591
Bus F2F Connections	2992
Total F2F Connections	28485
B2B Connection Pitch	5 um
Total B2B Connections	3024
DRAM Connection Pitch	25 um
DRAM Connections	3624

Figure 10.7.7: Die micrograph, floorplan, and technology data.

University Michigan, ISSCC and The Register



Dean Lewis, another GIT researcher, showed 3-D MAPS, a massively parallel processor using 64 custom cores stacked with a block 256 kilobytes of scratch pad memory. (Paper 10.6)

benchmark	memory bandwidth	measured power
AES encryption	49.5 GB/s	4.032 W
edge detection	15.6 GB/s	3.768 W
histogram	30.3 GB/s	3.588 W
k-means clustering	40.6 GB/s	4.014 W
matrix multiplication	13.8 GB/s	3.789 W
median filter	63.8 GB/s	4.007 W
motion estimation	24.1 GB/s	3.830 W
string search	8.9 GB/s	3.876 W

Figure 10.6.5: Bandwidth and power measurement results.

As might be expected, different tasks performed on 3D-MAPS have different performance profiles

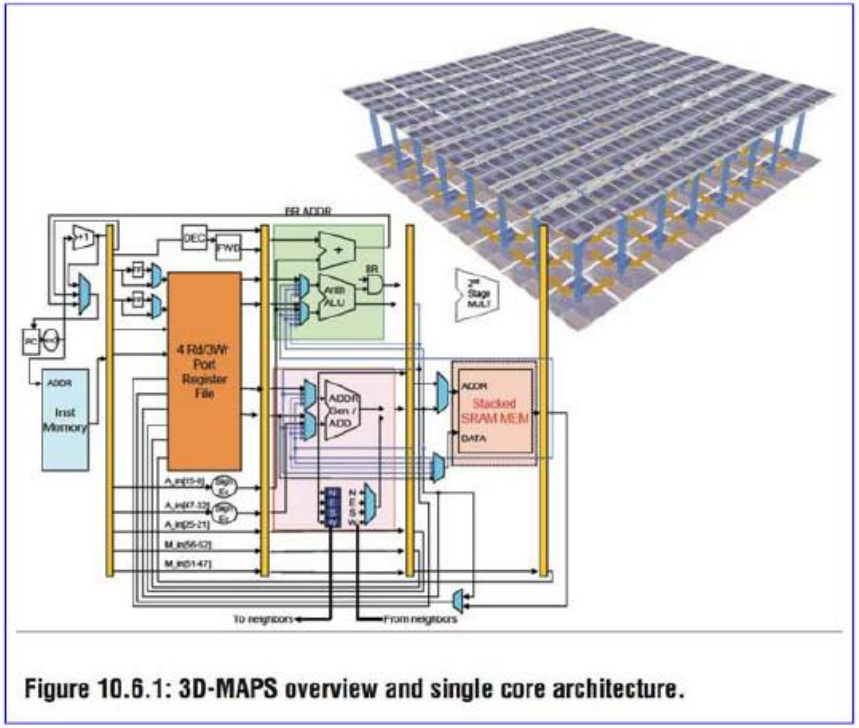
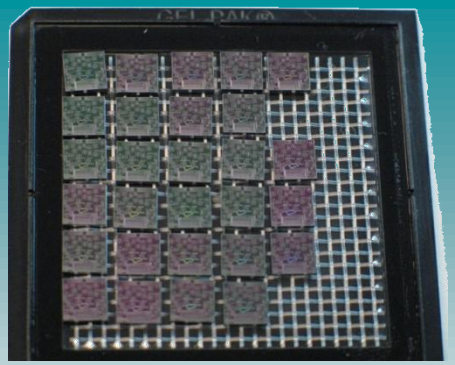
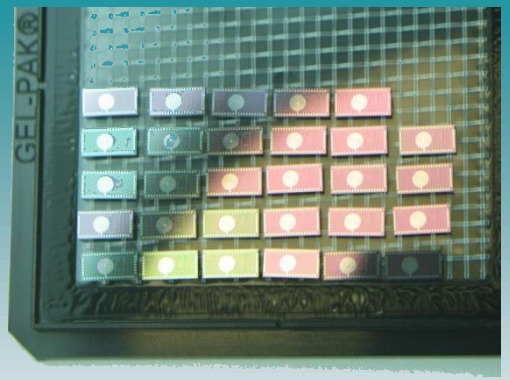
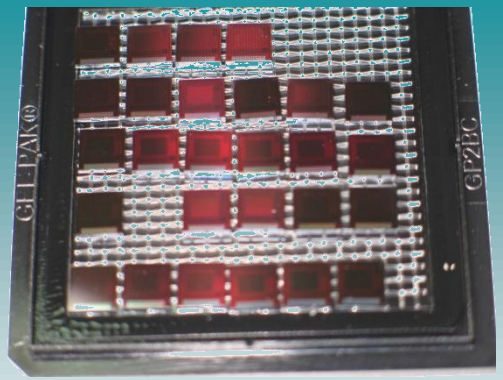
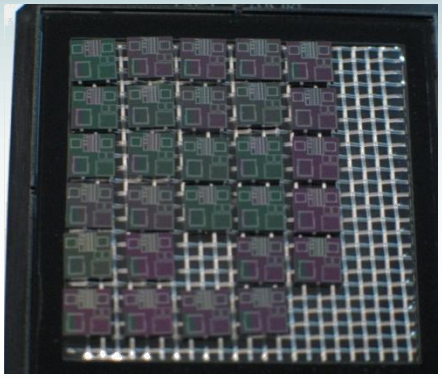
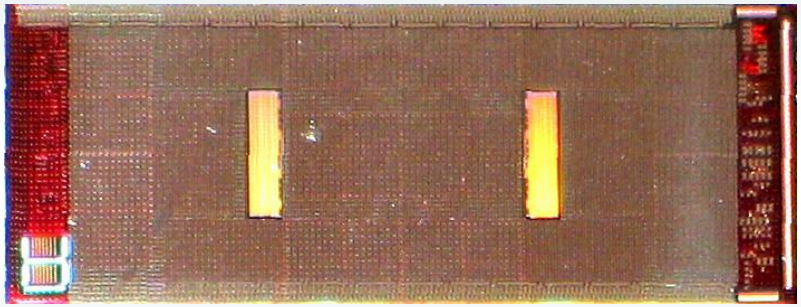
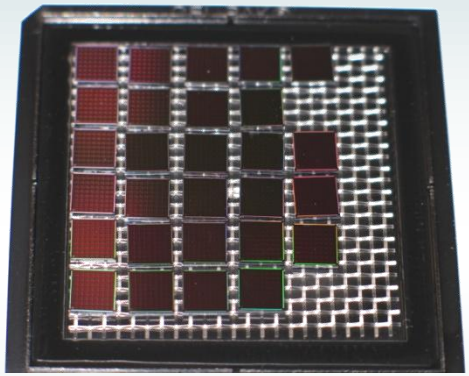
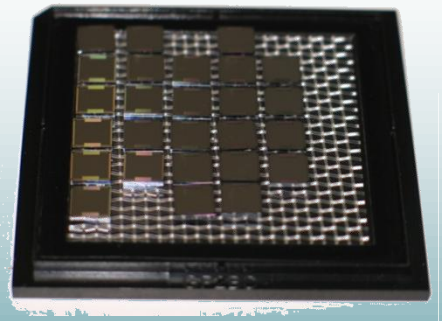
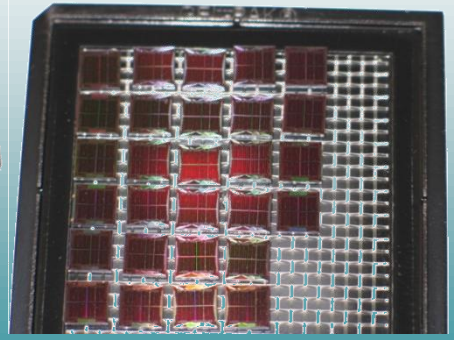
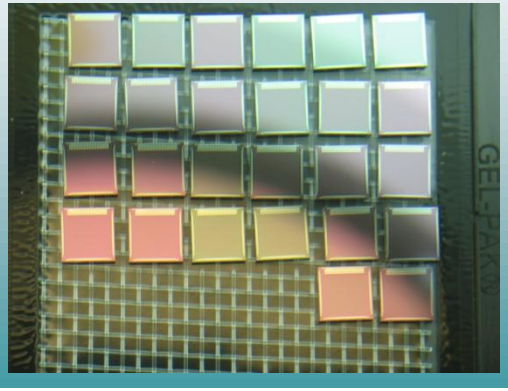
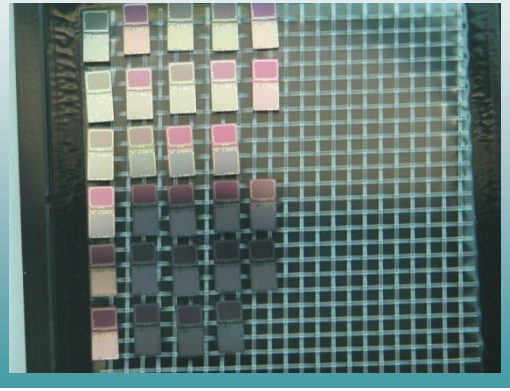


Figure 10.6.1: 3D-MAPS overview and single core architecture.

3D-MAPS's cores and layout may be simple, but version two is already on the way (click to enlarge) [8]

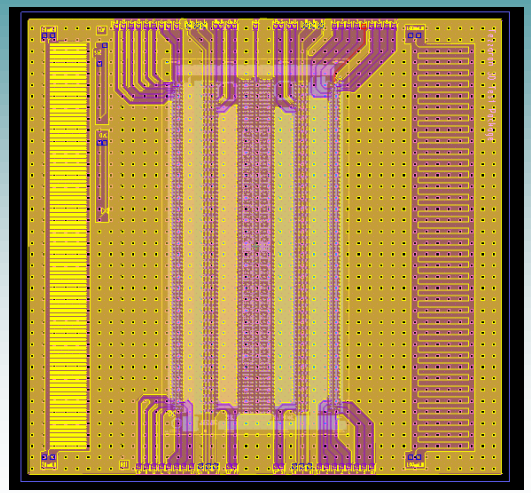
Georgia Institute of Technology, ISSCC and The Register

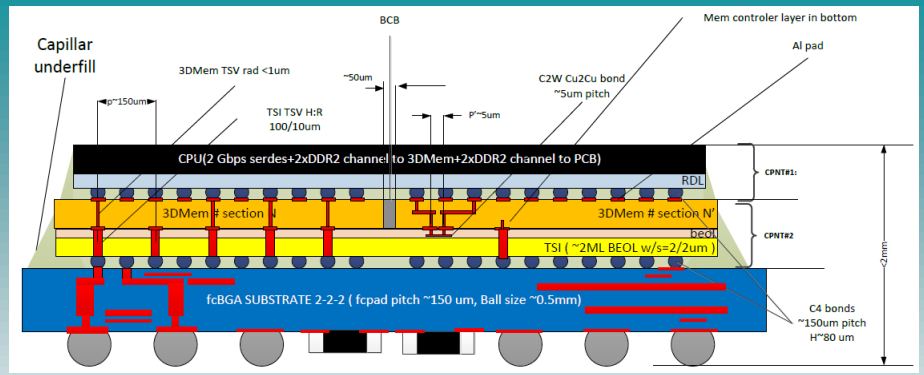











Tezzaron 3D Devices June/July 2011

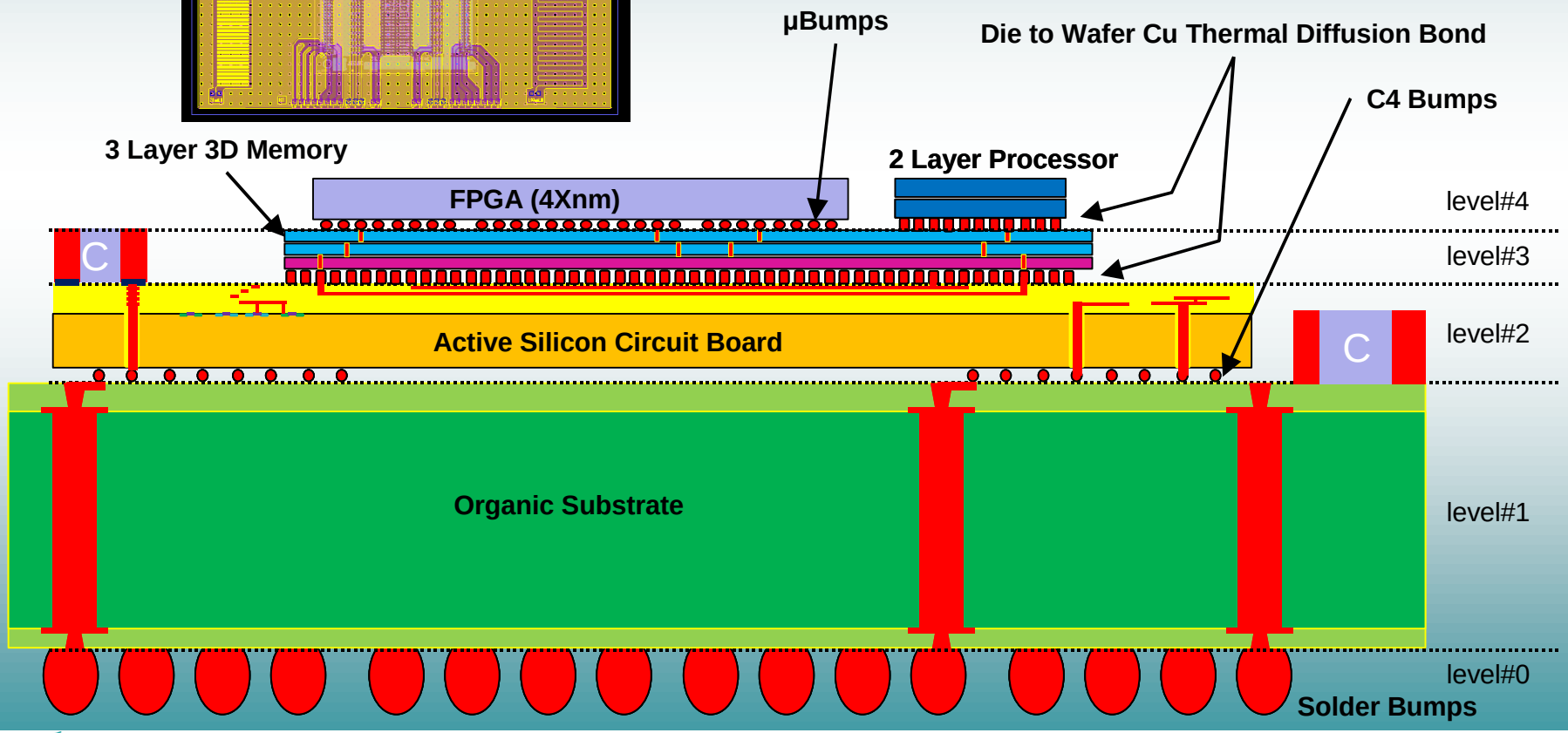
2.5/3D Circuits

IME A-Star / Tezzaron Collaboration



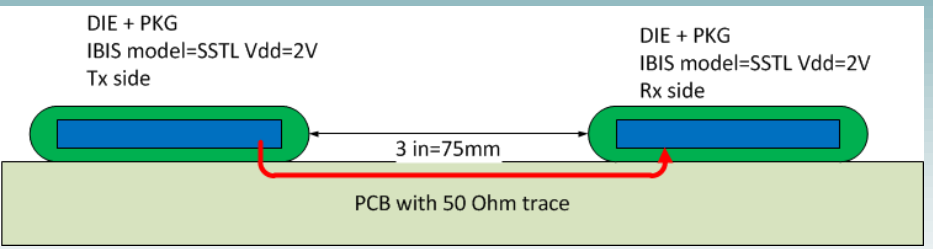


IME A-Star / Tezzaron Collaboration

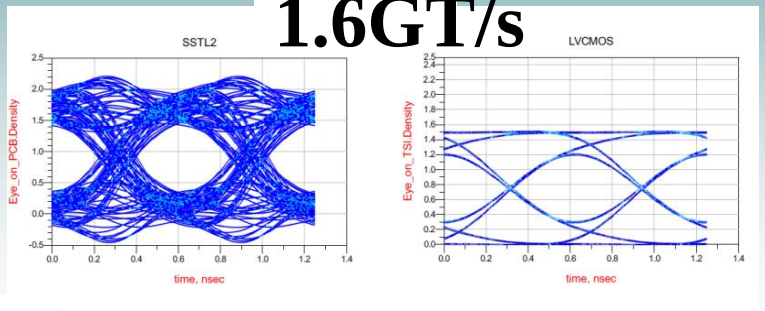


System in PCB vs in TSI (Courtesy A*STAR Institute of Microelectronics)

a)For system on PCB (each die is packaged),SSTL type of circuit are used (50 Ohm landscape design on PCB)
b) For system on TSI LVCMOS driver are used (BEOL wire design)

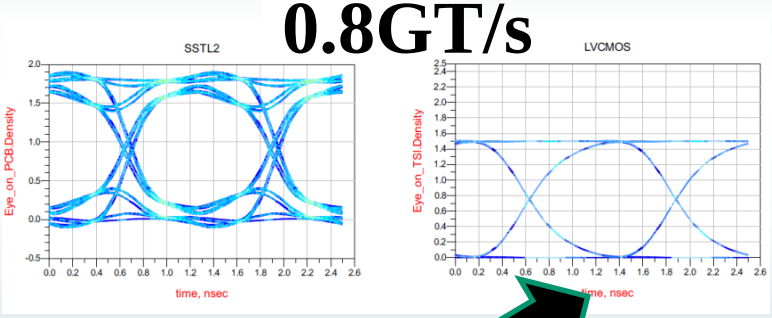


1.6GT/s



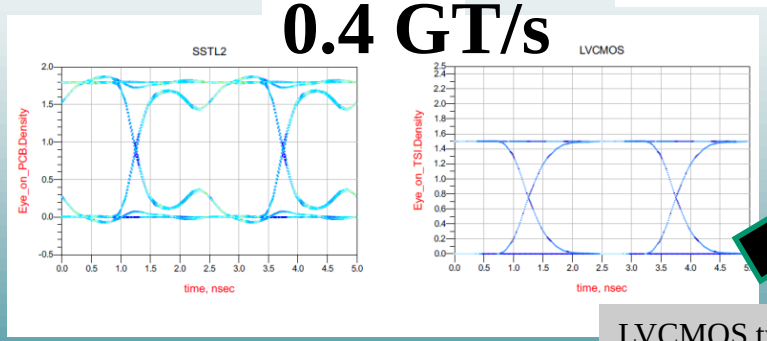
SSTL type of eye designed with 50 Ohm "landscape" (75 mm long) has increased SI penalty with Mem speed improvement

0.8GT/s

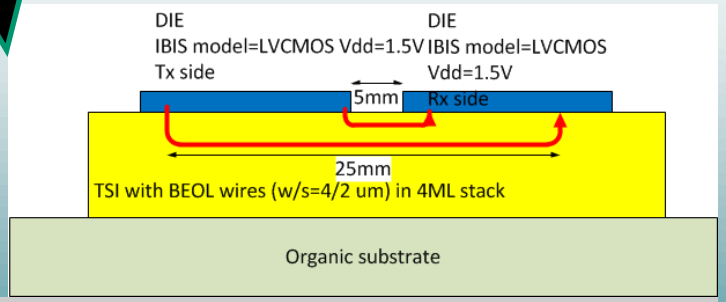


Agency for Science, Technology and Research
SINGAPORE

0.4 GT/s



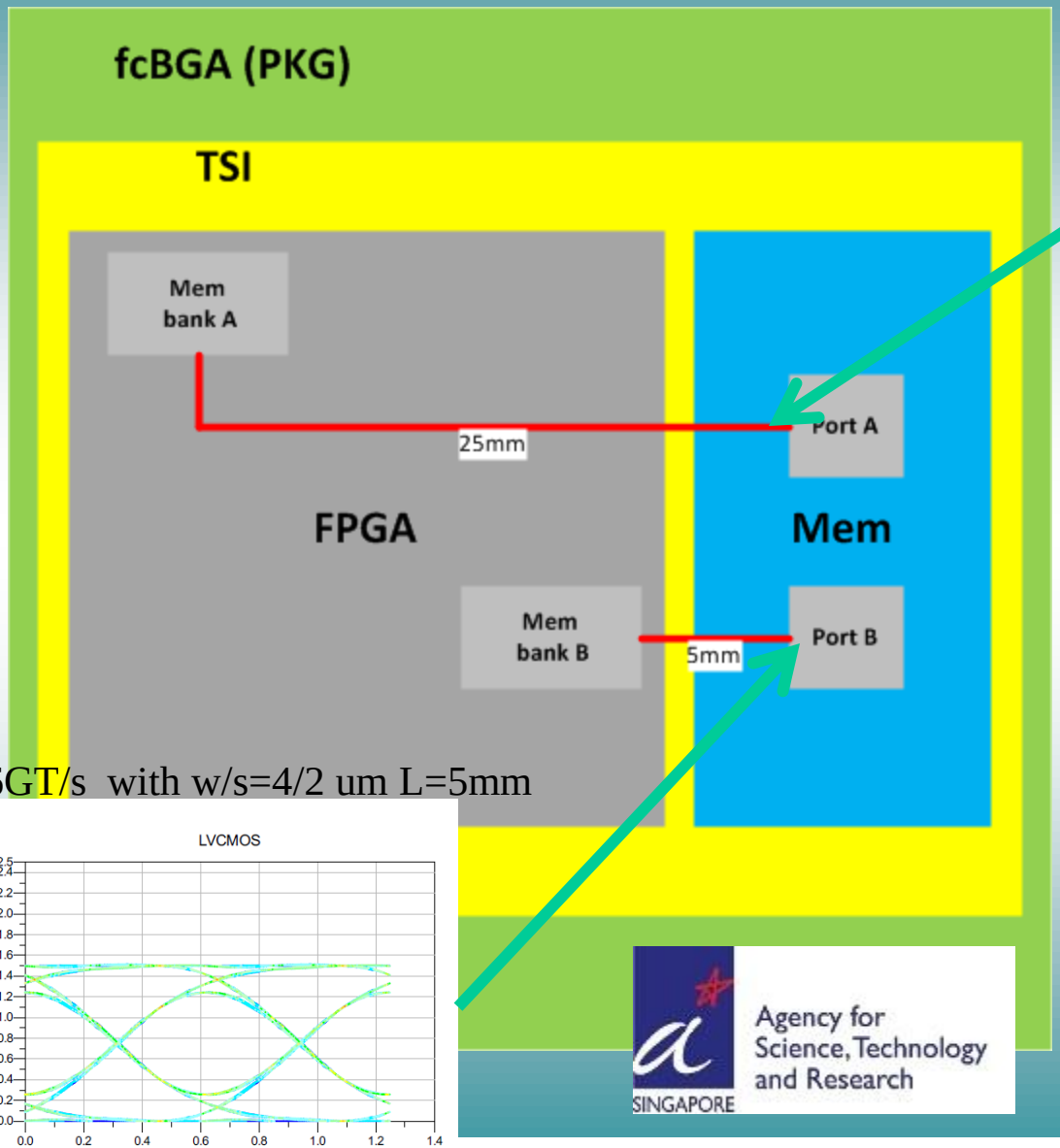
DDR, DDR2,DDR3



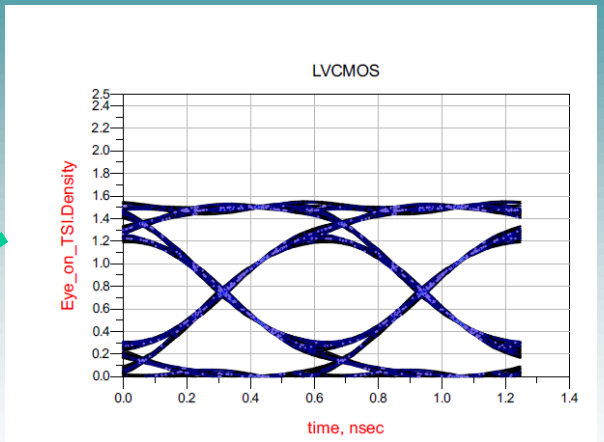
LVCMOS type of eye on BEOL wire (e.g w/s=10/10um) with 25 mm length,TSI "landscape" are more SI robust to Mem. speed improvement…..**showing margin for design in the eye**

Note : R,L,C for PKG is 0.3Ω,3.7nH,2.2pF

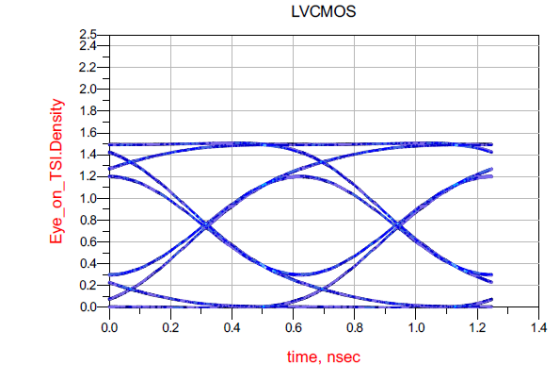
Layout considerations (Courtesy A*STAR Institute of Microelectronics)



1.6GT/s with w/s=4/2 um L=25mm



1.6GT/s with w/s=10/10 um L=25mm

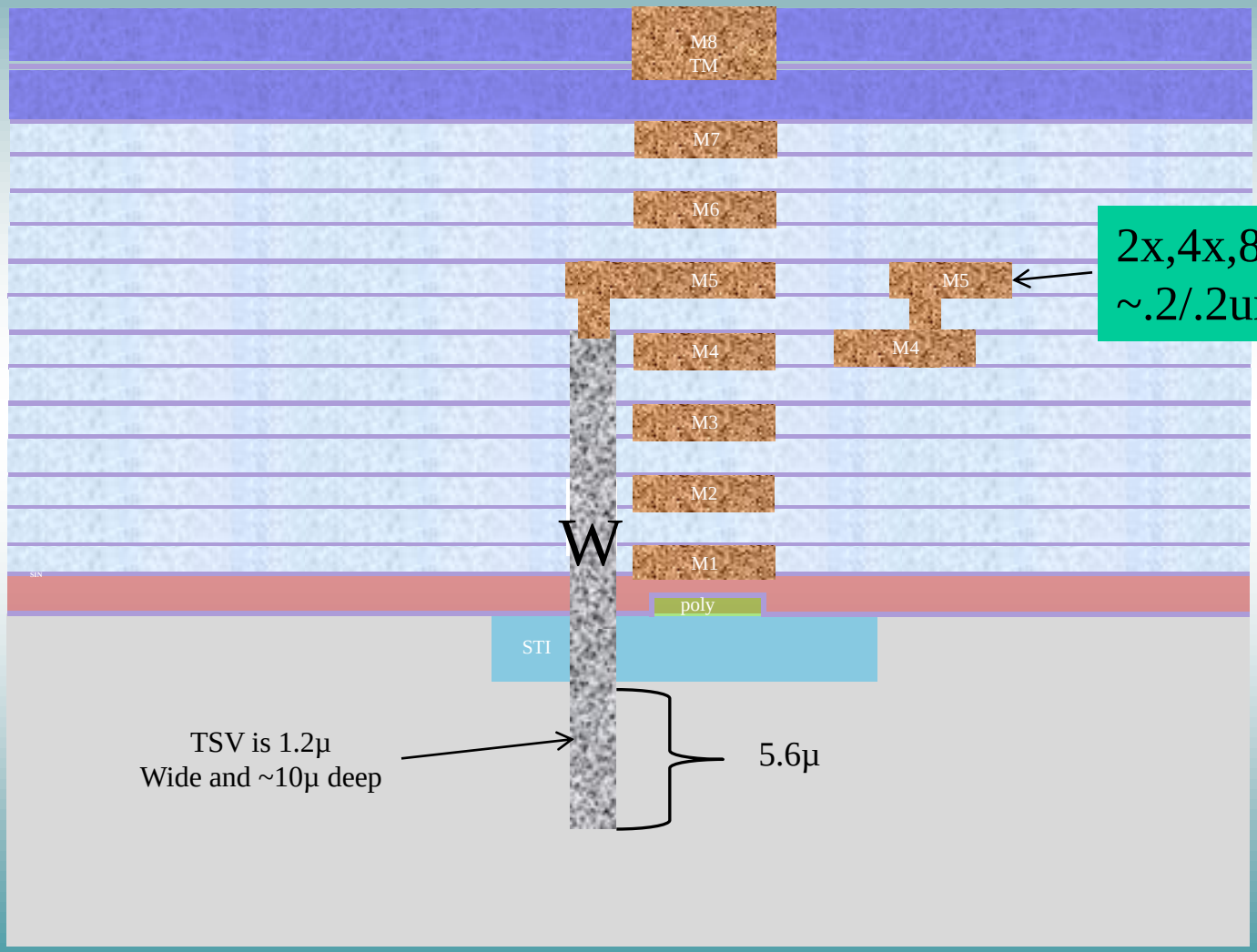


1.6GT/s with w/s=4/2 um L=5mm

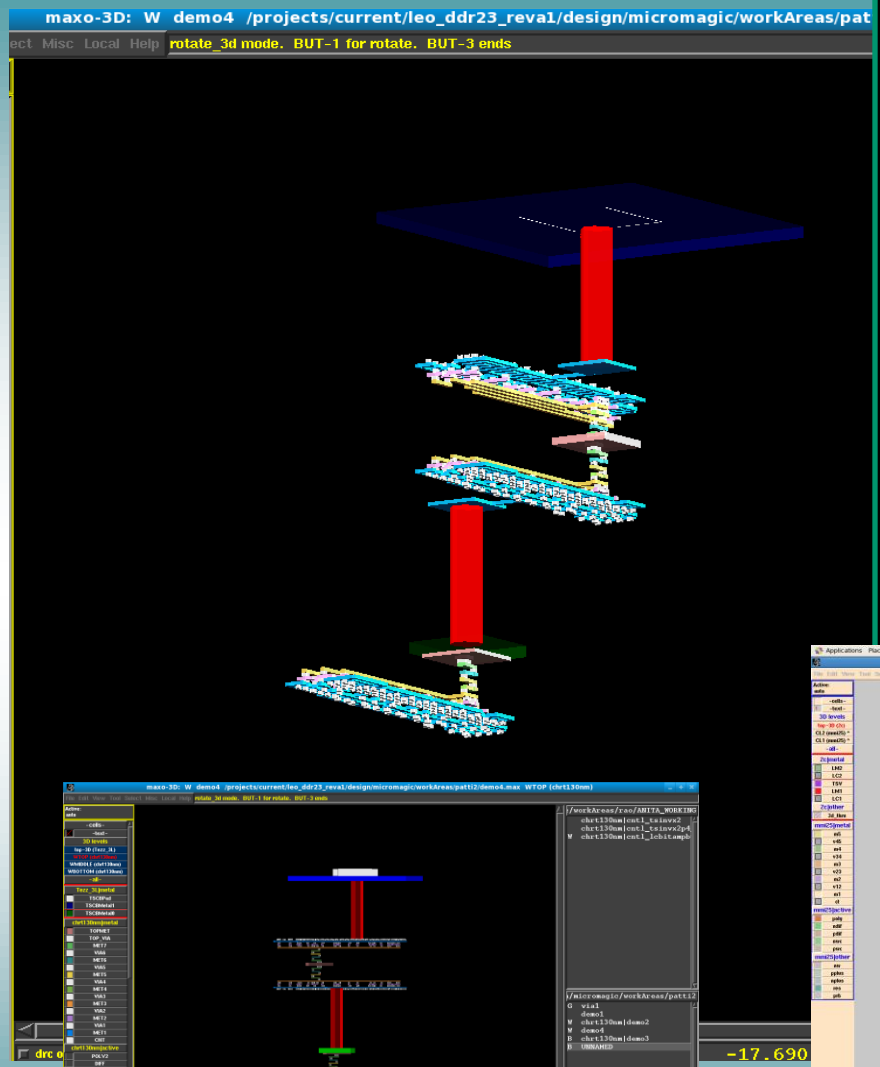
- RC penalty type is reduced on long line by using large wires(R drop) and large spacing (C drop)
- Byte lanes [8 wires of DQ and 1 wire of DQS (strobe)] with spacing S between wires S=4*(W+S) to eliminate crosstalk from parallel wires (2nd effect order)

Agency for Science, Technology and Research SINGAPORE

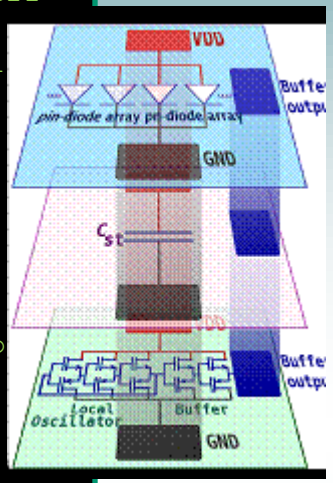
Near End-of-Line

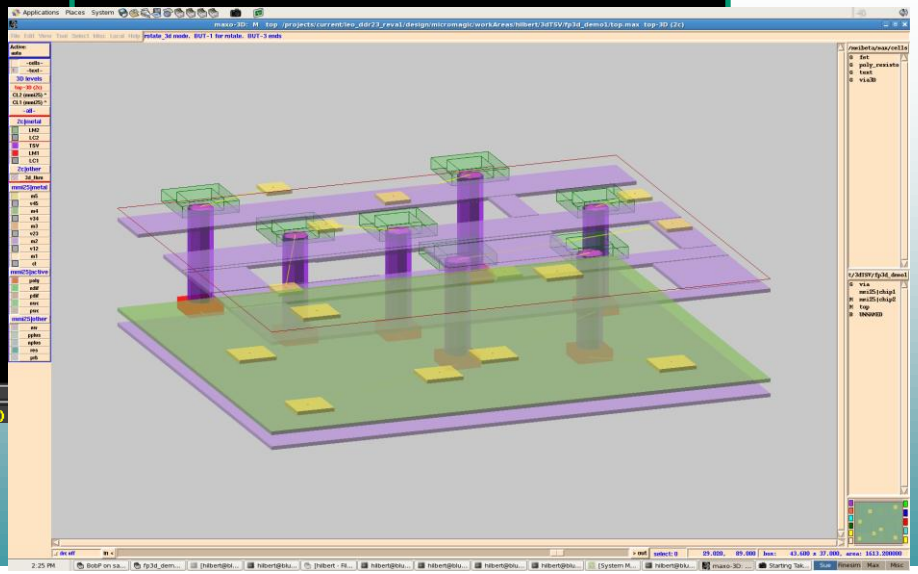


3D EDA Tool Support



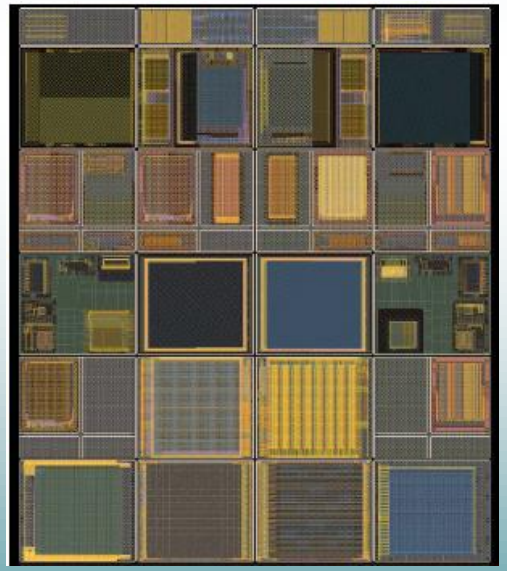

```
# 3D LVS Tech file
WAFER: 1
   LAYOUT TOP BLOCK: lvslayer1_1
   SCHEMATIC TOP BLOCK: lvslayer1
   GDSII FILE: lvslayer1_1.gds
   SCHEMATIC NETLIST: lvslayer1
   INTERFACE UP METAL: 1;0
   INTERFACE UP TEXT: 1;101
...
INTERFACE:
   LAYOUT TOP BLOCK: lvstop
   SCHEMATIC TOP BLOCK: lvstop
   GDSII FILE: lvstop_ALL.gds
   SCHEMATIC NETLIST: lvstop.sp
   BOND OUT METAL: 5;0
   BOND OUT TEXT: 5;101
```



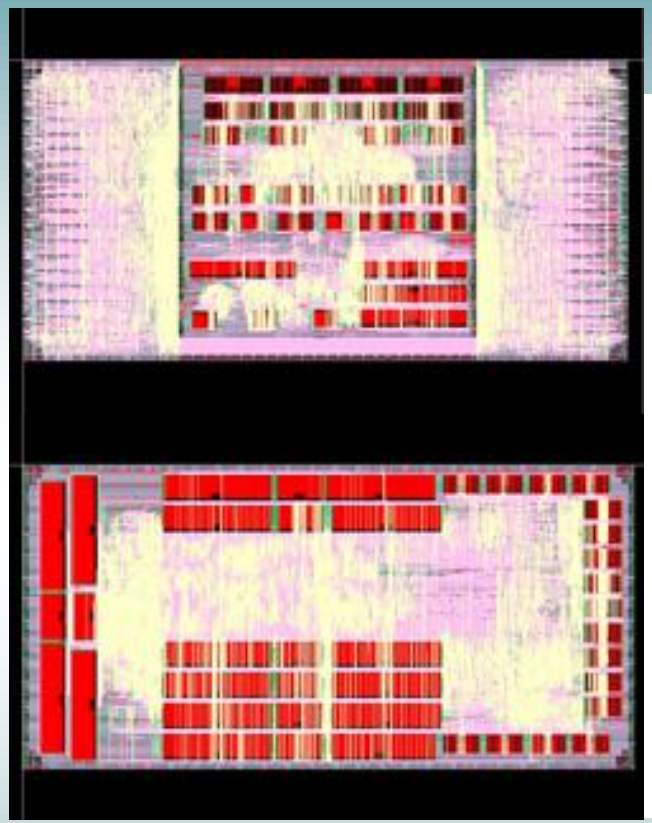


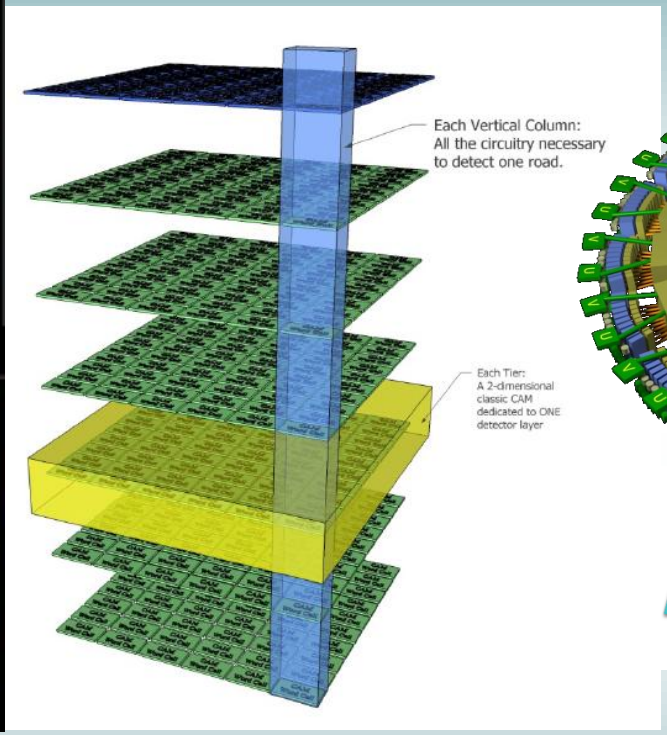
3D Opportunities

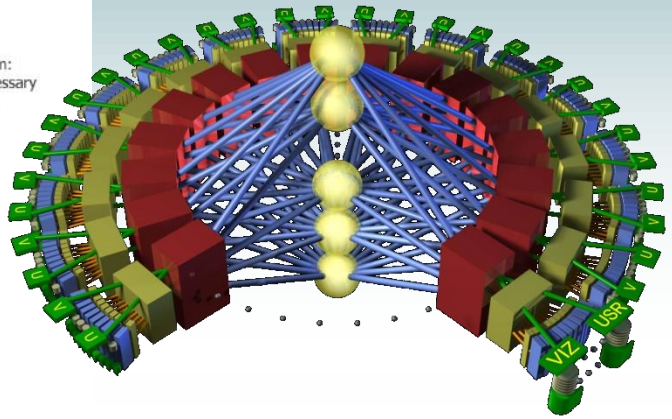
- Complete 3D PDK 8th Release
  - GF 130nm
  - Calibre, Synopsys, Hspice, Cadence MicroMagic 3D physical editor
  - Magma 3D DRC/LVS
  - Artisan standard cell libraries
- MOSIS, CMP, and CMC MPW support
  - 130nm, coming soon 65nm
  - Silicon Workbench
- Honeywell 150nm SOI
- 90nm NEOL TSV
- 40nm 3D logic
- IME silicon interposers, photonics
- eSilicon 3D solutions, organic interposers
- >100 devices in process
- >400 users

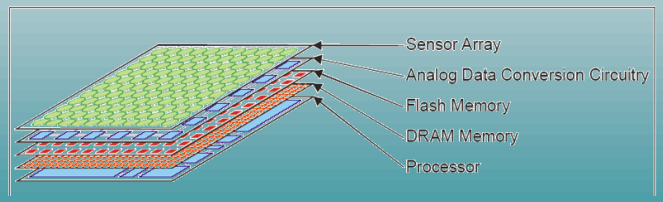


New Apps – New Architectures









Summary

- 3D Integration offers vast new opportunities!
  - New design approaches
  - New ways of thinking
  - Best of class integration of
    - Memory
    - Logic
    - RF
    - MEMS

Sensors

Computing

MEMS

Communications