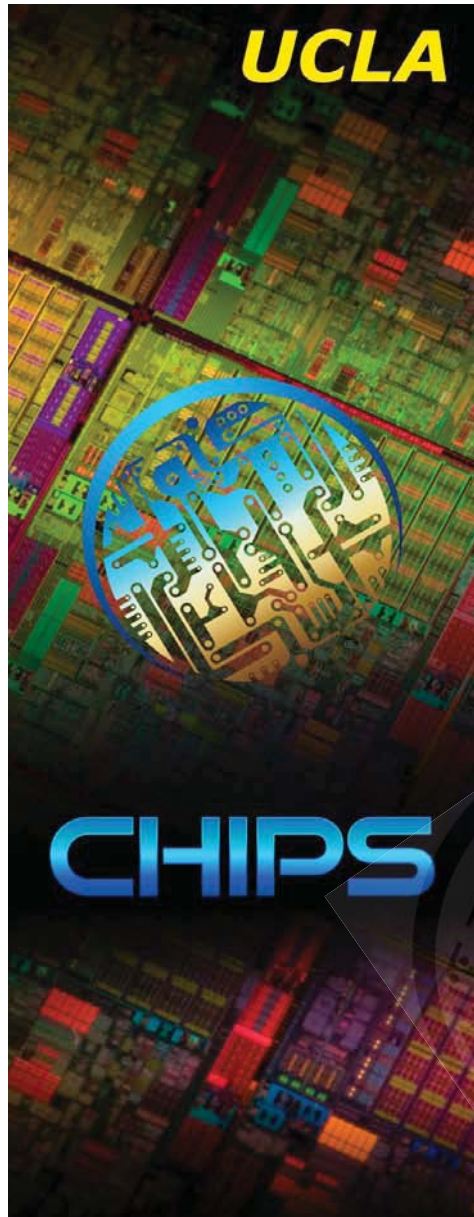


UCLA



CHIPS

CHIPS

Subu Iyer (s.s.iyer@ucla.edu)

UCLA ENGINEERING
Henry Samueli School of
Engineering and Applied Science
Birthplace of the Internet



CHIPS
CENTER FOR HETEROGENEOUS INTEGRATION
AND PERFORMANCE SCALING

Three questions

- What are we trying to do ?
- Why are we doing this ?
- How are we planning to do this ?

What are we trying to do ?

Develop an “app-like” environment for Hardware that can

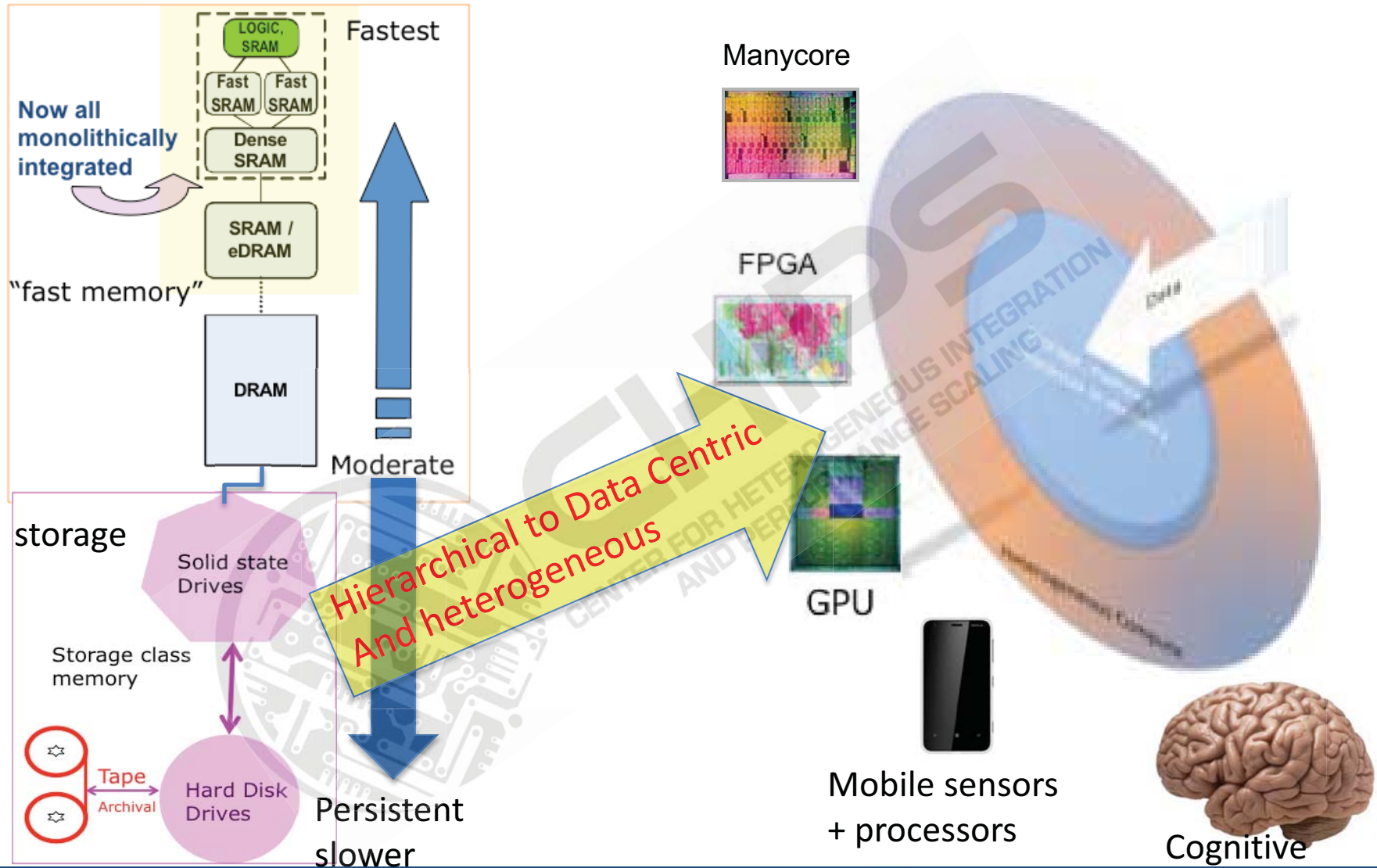
- Cut the time to market by 5-10X
- Cut the NRE cost by 10-20X
- Allow extreme heterogeneity including extensions to cyber-physical systems
- Develop a sophisticated manufacturing workforce



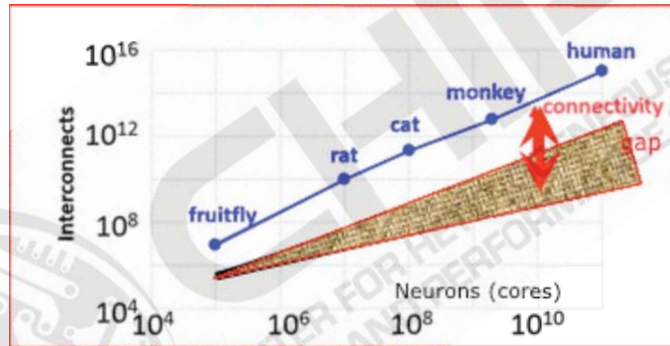
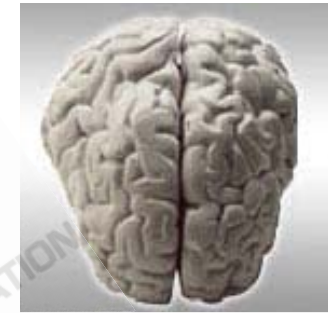
Why now ?

- The application space is transforming
- CMOS scaling is saturating
- **Systems are getting more heterogeneous and complex**
- Holistic system Integration is the next frontier of Moore's law

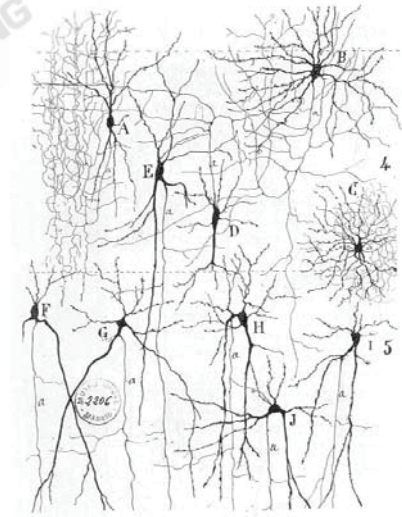
Computing is transforming



“cognitive” computing – brawn Vs. brain



It is unlikely that we can just scale our selves out of the huge gap (10^3 to 10^4) in power and interconnect density



The Power of System on a Chip (SoC)

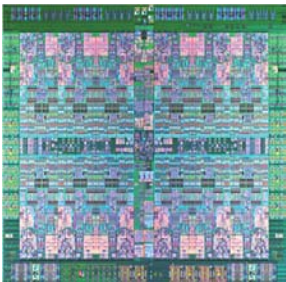
Transistor scaling has made this possible



1964 - Transistor

SLT module

6 transistors, 4 resistors



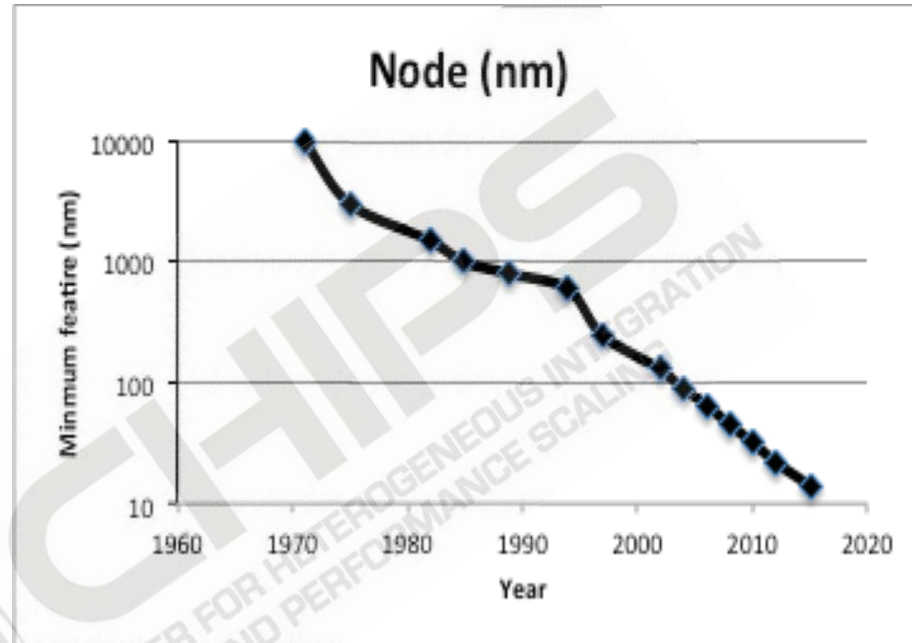
2014 – POWER8 Processor

22nm SOI eDRAM technology, 650mm²

12 cores and 96MB of on-chip memory

4.2 billion transistors

Courtesy IBM



But as SoCs have gotten more complex

- NRE costs have skyrocketed
- Time to market has become huge
- Manufacturing costs have grown
- and yields have plummeted

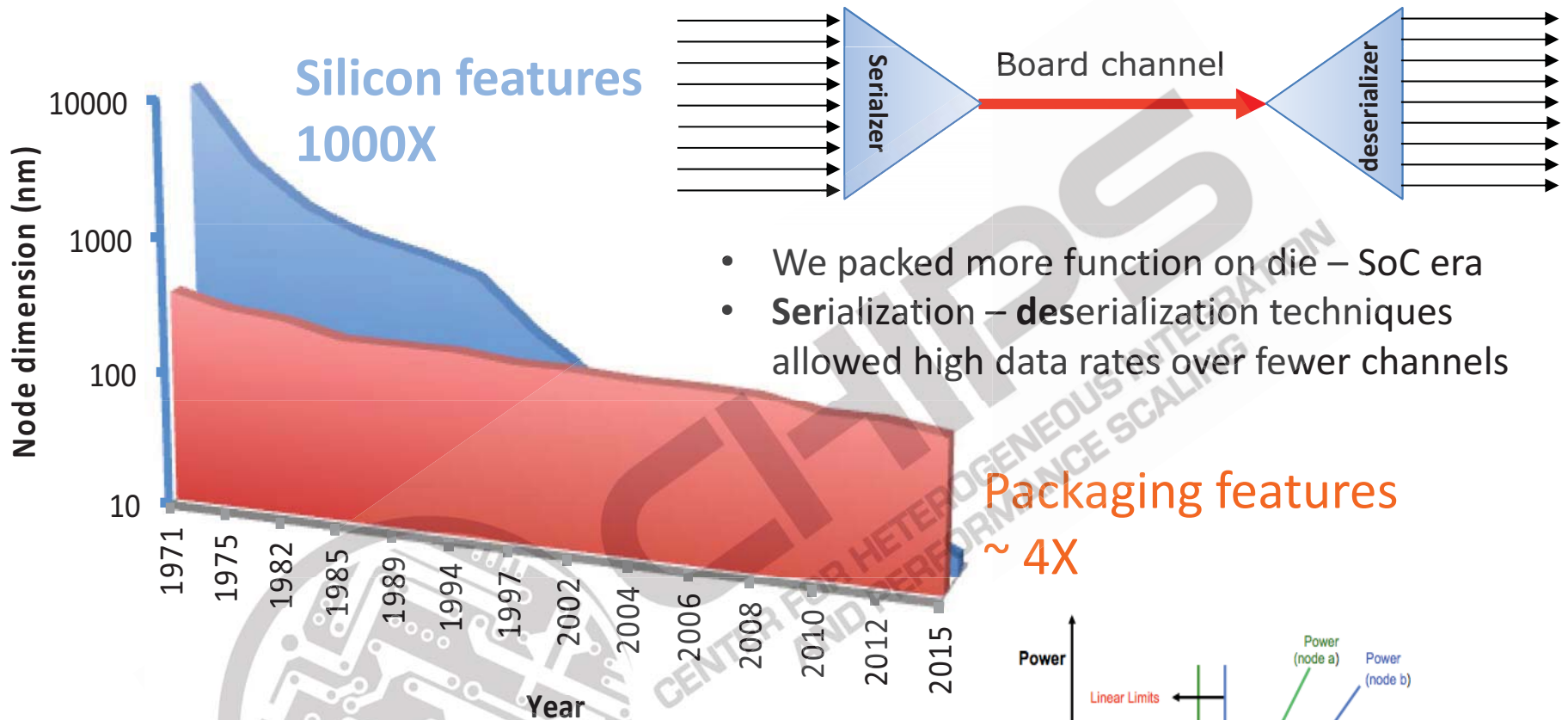
What would Yogi Say ?



1925-2015

So, if you can't scale the chip
you should try to scale something else !

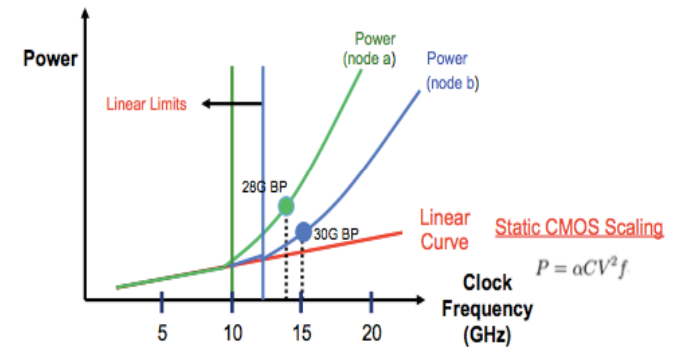
Package/Board Features have scaled modestly



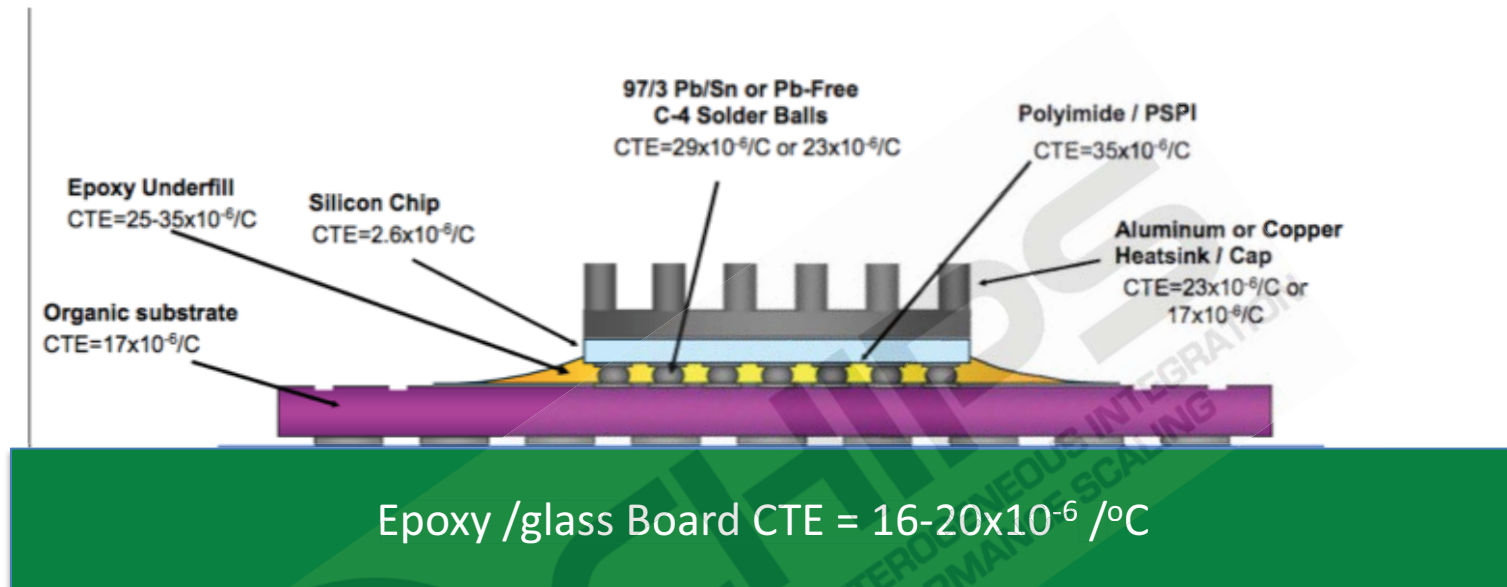
- We packed more function on die – SoC era
- **Serialization** – **deserialization** techniques allowed high data rates over fewer channels

Packaging features
~ 4X

- Boards are large
- “Serdes” area (>25%) and power (>25%) getting out of hand



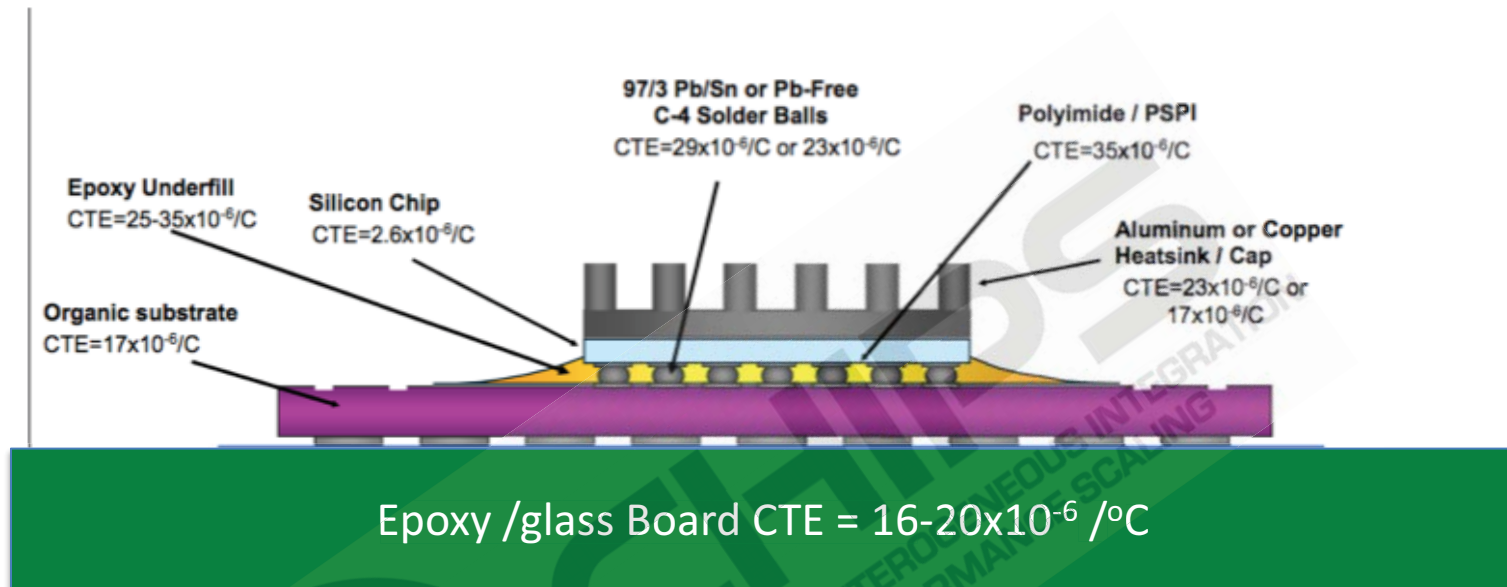
Do we need Packages ?



Packages are supposed to:

- Protect the chip – mechanically and thermally
- Connect the chip electrically to other chips
- Allows mission mode testing of the chip

Do we need Packages ? – No!



Packages are supposed to:

- Protect the chip – mechanically and thermally (not really)
- Connect the chip electrically to other chips (by progressive fan-out)
- Allows mission mode testing of the chip (this is true)

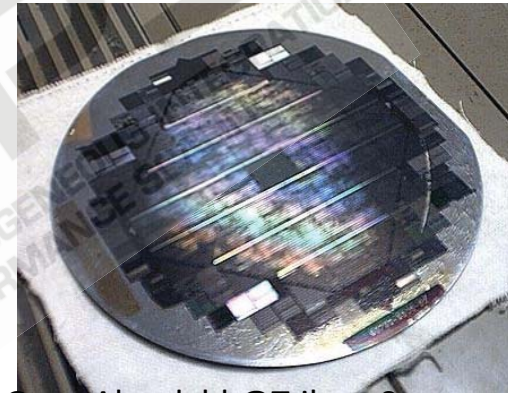
We can scale the package by getting rid of it

And mount the bare die directly on the board

Shrinking the board

- Without question the smallest board is a single large die with all the system functions on it

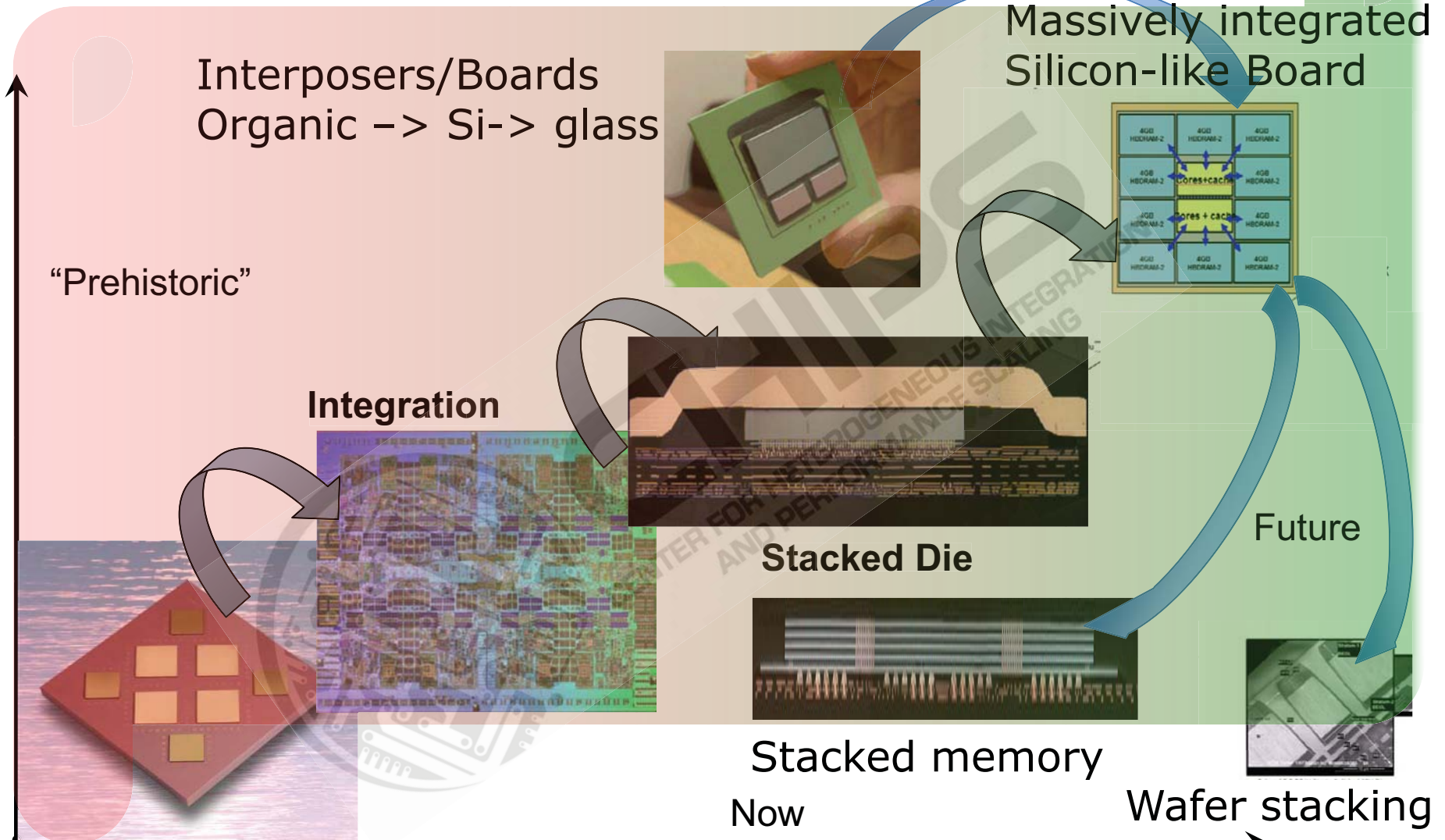
Amdahl eventually declared the idea would only work with a 99.99% yield, which wouldn't happen for 100 years.



Gene Ahmdahl @Trilogy Systems

- Hence the quest for the largest Yieldable Unit – the SoC (reticle size today will yield at 3-15%)

Evolution Of System Integration

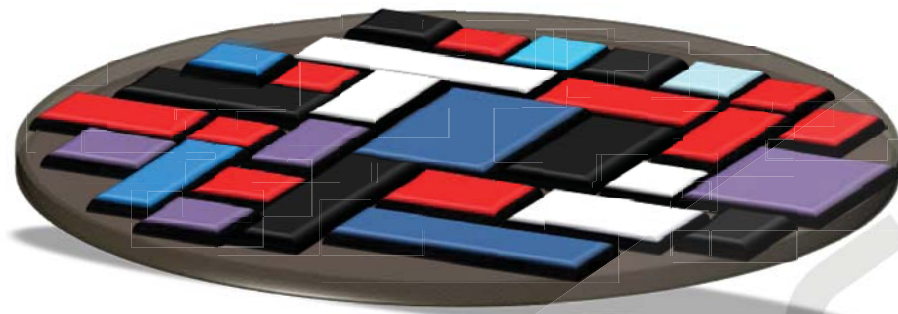


Mega SoFs by re-integration on an Interconnect Fabric

The "right" interconnect fabric

- Mechanically robust (flat, stiff, tough...)
- Capable of fine wiring, fine pitch interconnects
- Thermally conductive
- Can have active and passive built-in components

Silicon Fits the Bill in many cases

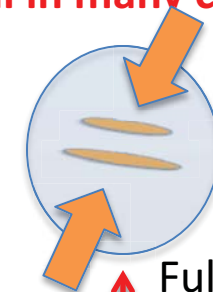
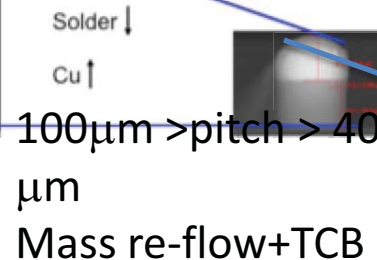


Challenges:

- Warpage
- Topography
- Assembly /Thru'put
- Thermal

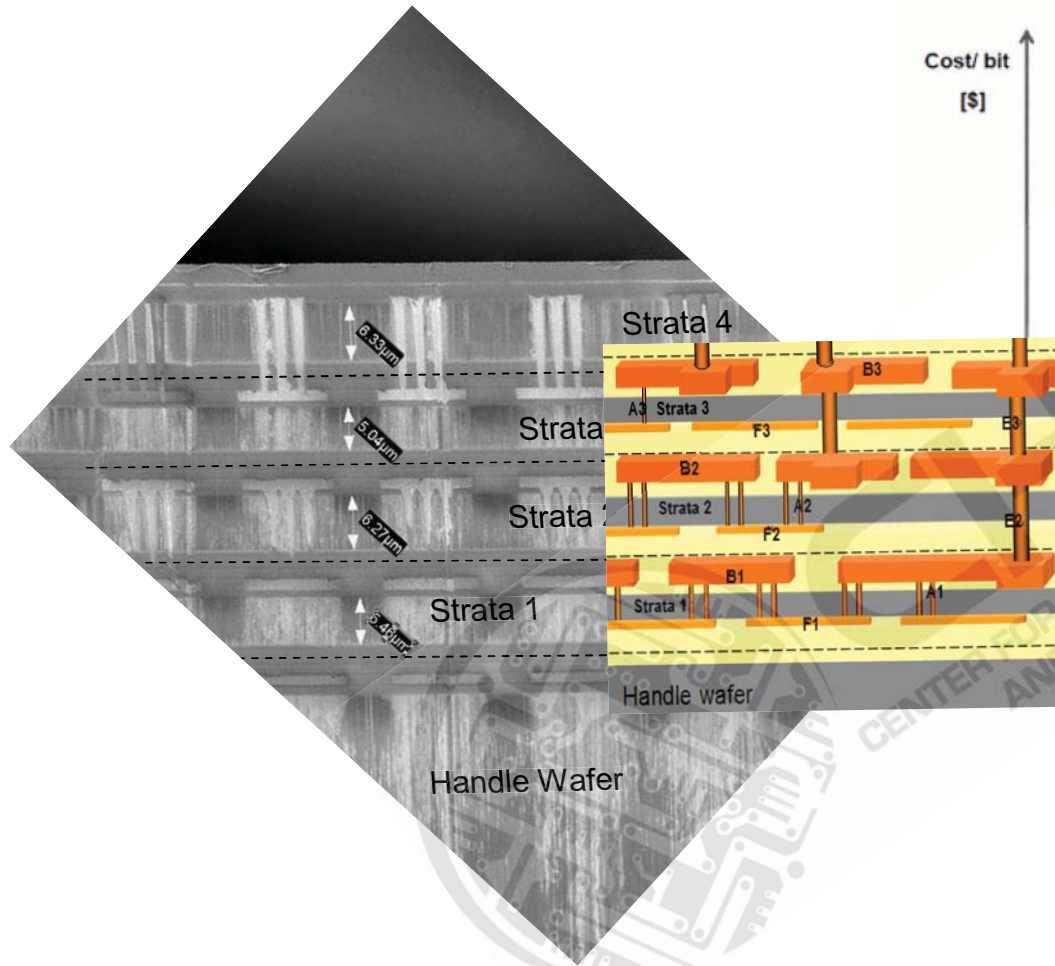


>100 μm pitch
Mass reflow

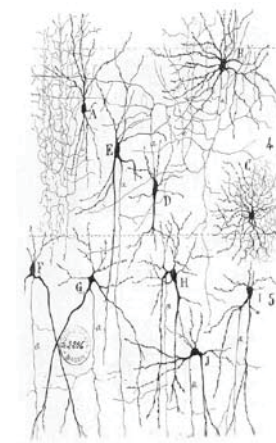
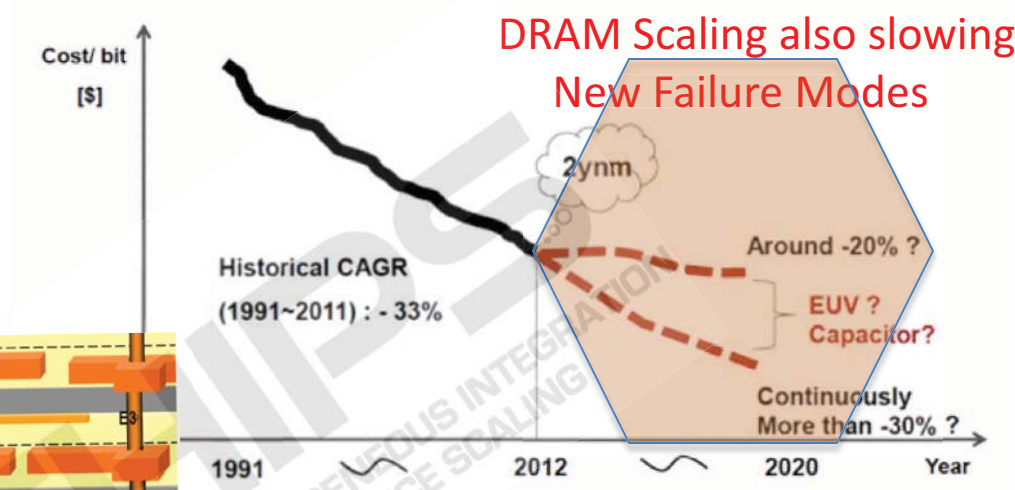


Full contact – TCB
Proximate
• Inductive
• Capacitive

Wafer to Wafer integration at tight pitch



4-layer wafer stack w/ 2.5 μm Pitch



Highly interconnected Microsystems

CHIPS Framework

Tier 1 Equipment Partners

New Tool Concepts
Tool Development
Scale-Up

Materials

Fine Pitch Interconnect
Substrate Materials
Warpage, Stress Flexible
Materials
Thermal solutions
Materials for Additive Mfg.
Reliability

Applications & Architecture

Heterogeneous Systems
Approximate Computing
Cognitive Computing
Fault Tolerance
Supply chain Integrity
Security
Memory Subsystems
Processing in Memory DFT
Network on Board

Integration

*A processing facility
for
Interconnect Fabric
(IF) & assembly*

- Silicon processing
- Glass
- Flexible substrates
- Additive manufacturing
- Thermal compression bonding
- Wafer thinning
- Wafer-wafer integration

Design Infrastructure

Thermo-Mechanical
Electrical
Tools
Partitioning
DFT
Active IF Design

Devices/Components

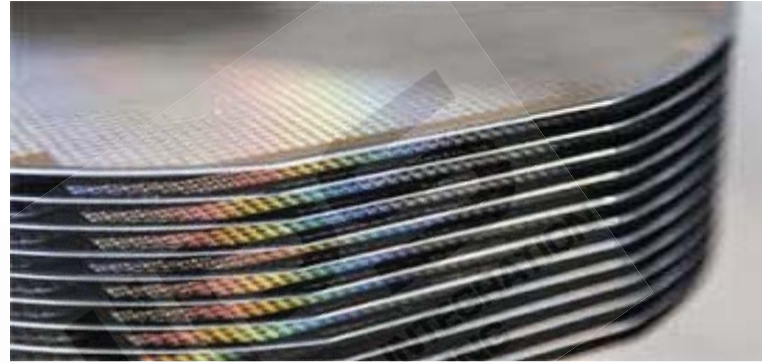
Novel switches
New memory
MEMS
Sensors
Passives, antennae
Medical devices

Tier 1 Foundry Partners

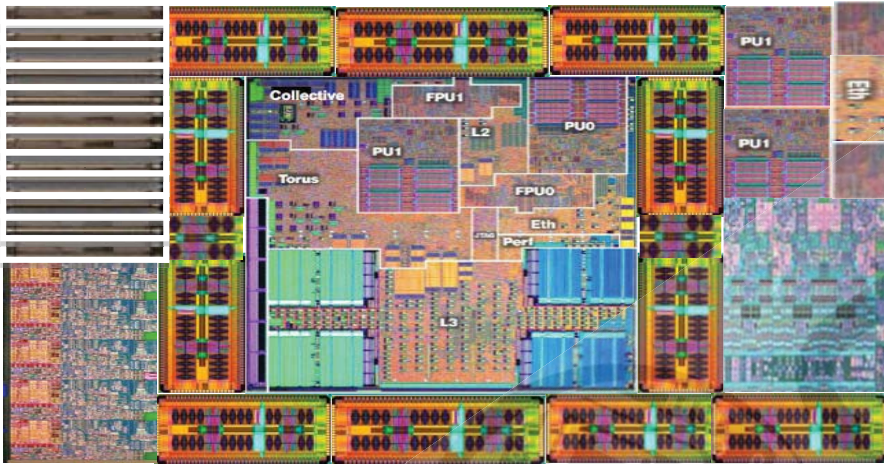
Si, Compound Semis,
MEMS, and OSATs



Test Beds



Desktop Datacenter



“Mega SoI^F*”
(~ wafer dimensions)

*System on Interconnect Fabric

CENTER FOR HETEROGENEOUS
AND PERFORMANCE SCALING

μ -Yodabot



autonomous origami μ -Yodabot that thinks,
walks, flies, swims and does stuff.....

Chip Assembly



UCLA

Summary

- We are in the midst of a significant hardware transformation
 - Semiconductor Scaling is saturating
 - Computing models are changing
 - Systems are getting more heterogeneous
 - SoCs design costs and Times-to-Market huge
- CHIPS will drive a much more holistic Moore's Law
- We can't do it alone, we have to do it together
 - Industry Partners
 - University Partners: GINTI (Tohoku U), Binghamton U, Georgia Tech,
 - Government Agencies

CHIPS

