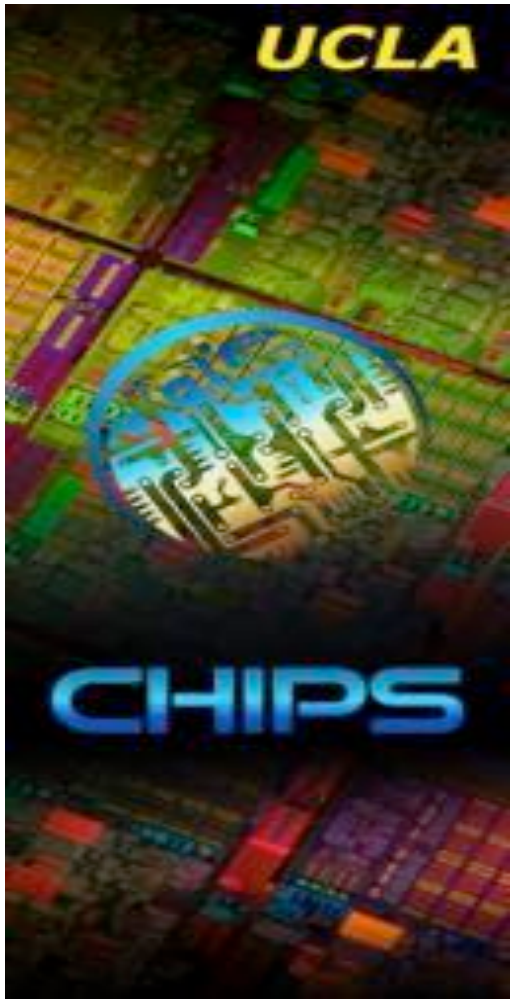



# Heterogeneous Integration, Dielets and Chiplets

## – why the hype ?

Subramanian S. Iyer
(s.s.iyer@ucla.edu)
chips.ucla.edu

Center for Heterogeneous Integration and Performance Scaling

UCLA Samueli School of Engineering

iMAPS GBF March 4, 2020

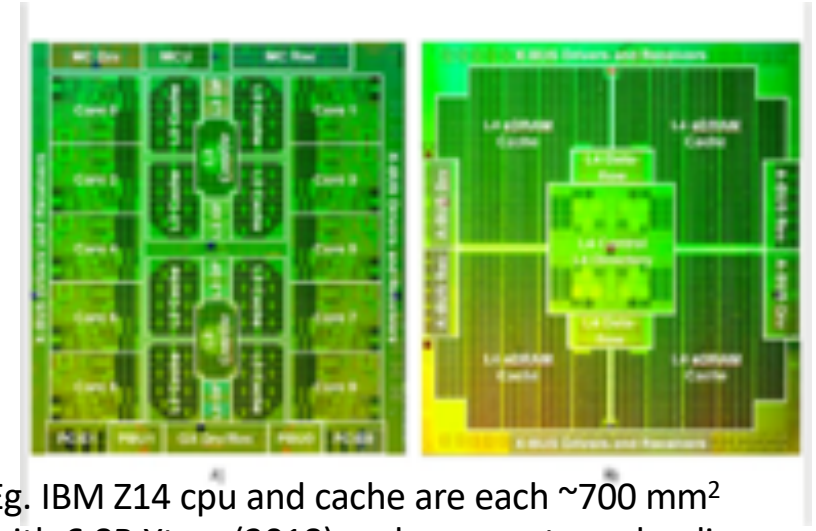
CHIPS CENTER FOR HETEROGENEOUS INTEGRATION AND PERFORMANCE SCALING

# Some reading material:

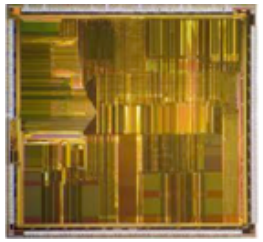
1. Spectrum article (October issue)
2. IBM J of R&D article (early acess)


# Some observations

- If Moore's law has enabled miniaturization, why have chips gotten larger ?
  - More complex problems
  - More cores @ higher clock speeds
  - More cache memory
- Main memory capacity and access limits performance
- Power density challenges - more "dark" silicon
- I/Os take up more space and power as system size increases

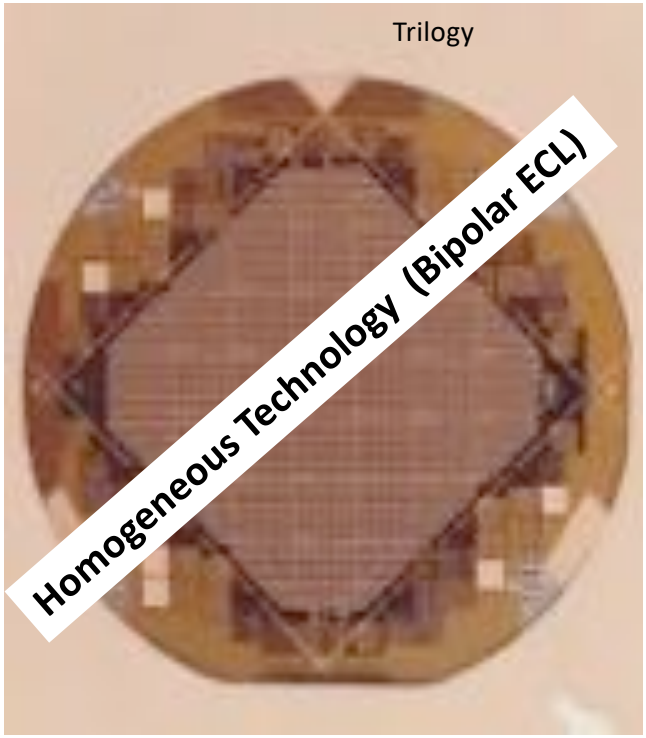


Eg. IBM Z14 cpu and cache are each ~700 $mm^2$ with 6.8B Xtors (2018) and a separate cache die

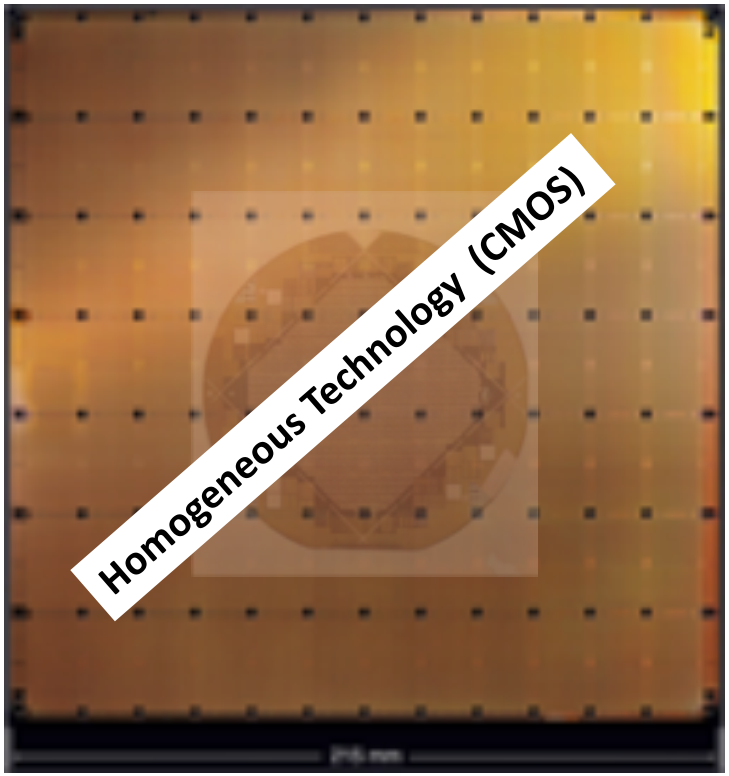


Intel Pemtium cpu ~300$mm^2$ -3.1 Million Xtors (1993)

# Wafer Scale Integration then and now

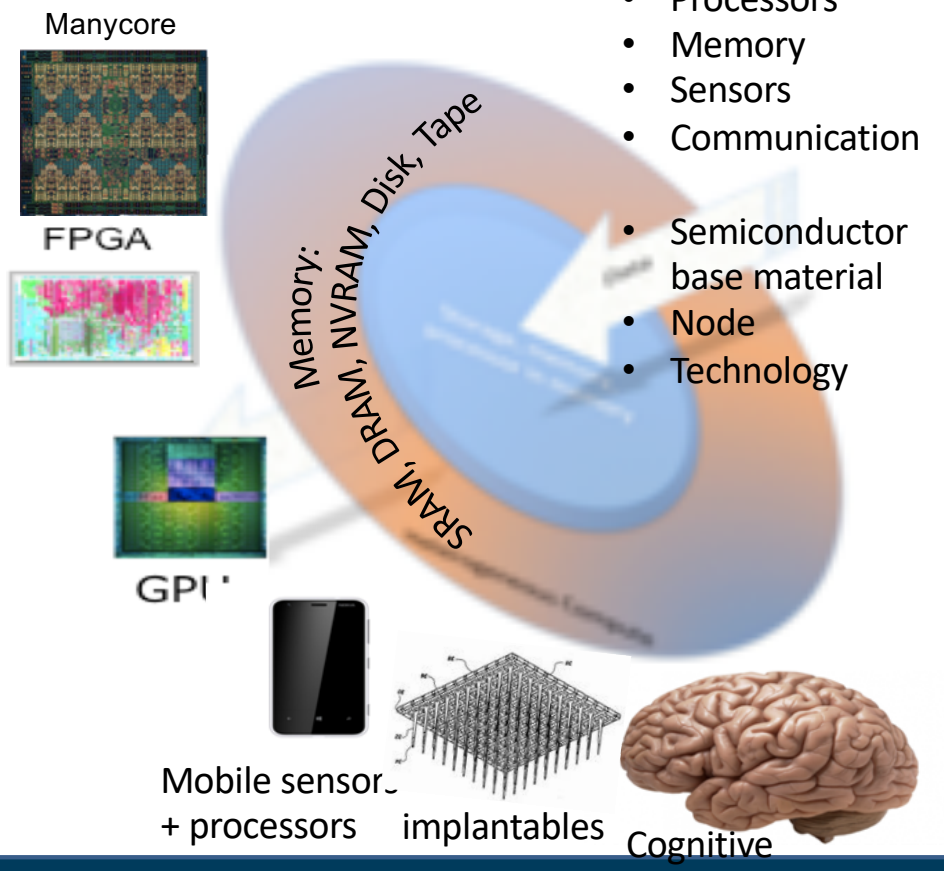



Bumped 100 mm wafer (Ca 1982)




Cerebras (2019)

# Heterogeneity, Performance and Scaling




- Processors
- Memory
- Sensors
- Communication

- Semiconductor base material
- Node
- Technology

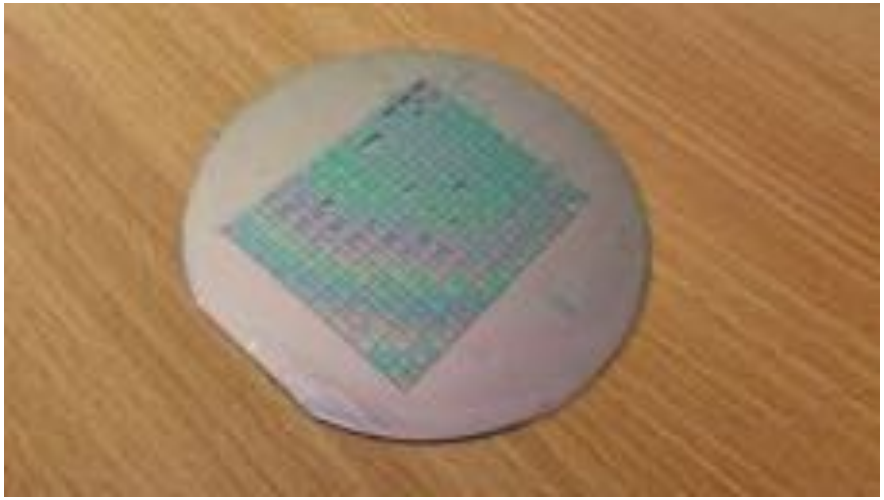
- For the Raw Device: scaling = performance
- For the interconnect: Not straightforward
- For System Functionality: Heterogeneity = performance
- For Cost/Function: Heterogeneity
- For Time to Market: Heterogeneity
- For power/energy: scaling, interconnects, heterogeneity, and *architecture*

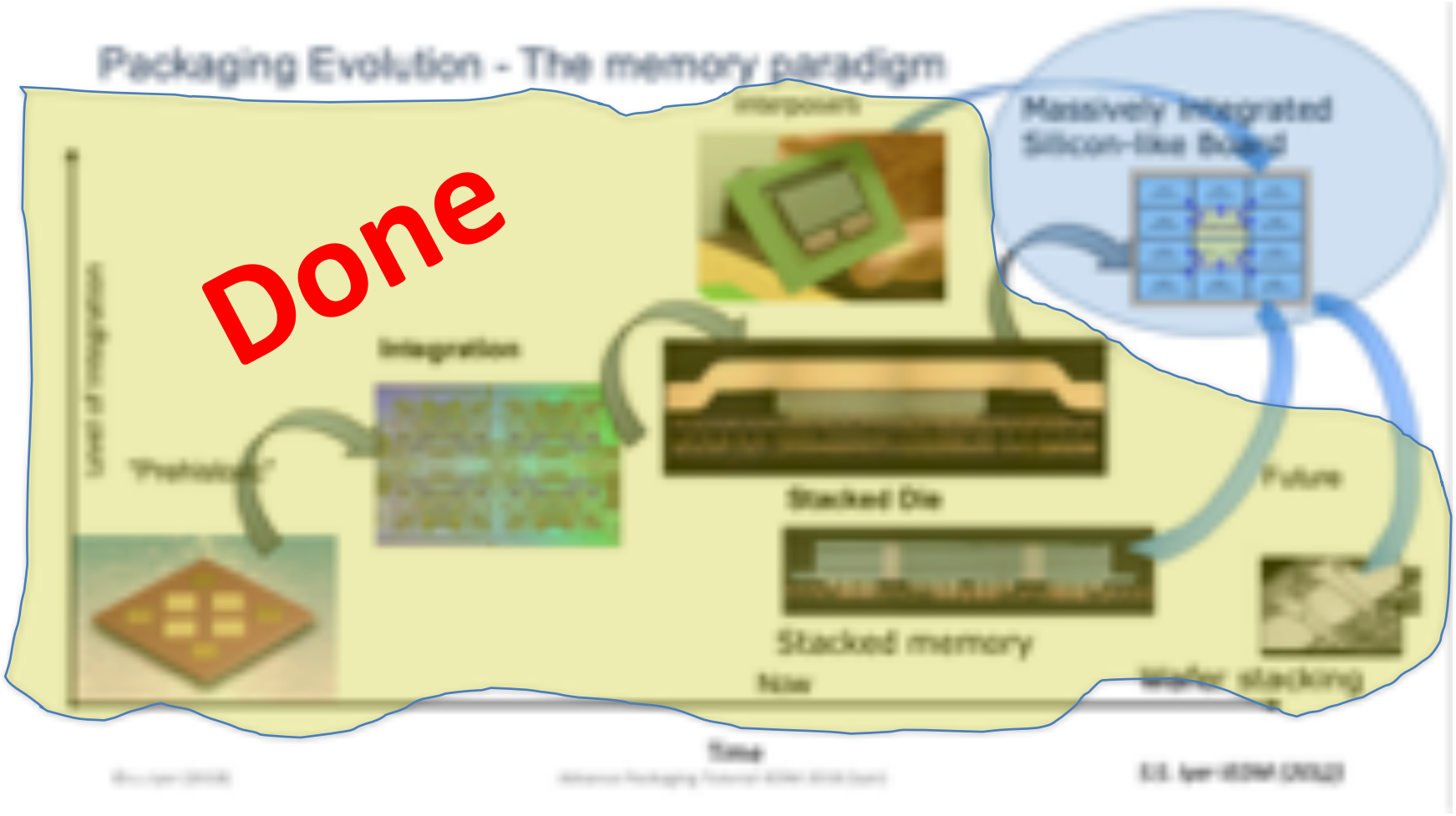
UCLA Samueli School of Engineering

CHIPS CENTER FOR HETEROGENEOUS INTEGRATION AND PERFORMANCE SCALING

# Wafer Scale Integration @UCLA CHIPS

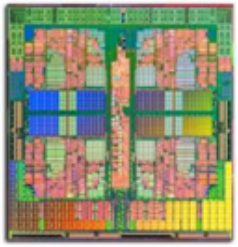
Lets revisit this concept but
with a different approach



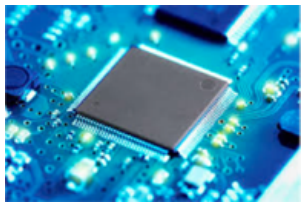

Packaging Evolution - The memory paradigm
Done
Interposers
Massively Integrated
Silicon-like Board
Integration
Stacked Die
Future
Stacked memory
Now
Wafer stacking
Time

# Chiplets or Dielets ?

A **die**, in the context of integrated circuits, is a small block of semiconducting material on which a given functional circuit is fabricated. We usually think of dies as bare.



A chip is more or less the same thing but connotes the design aspects rather than the physical attributes. We usually think of chips as packaged.



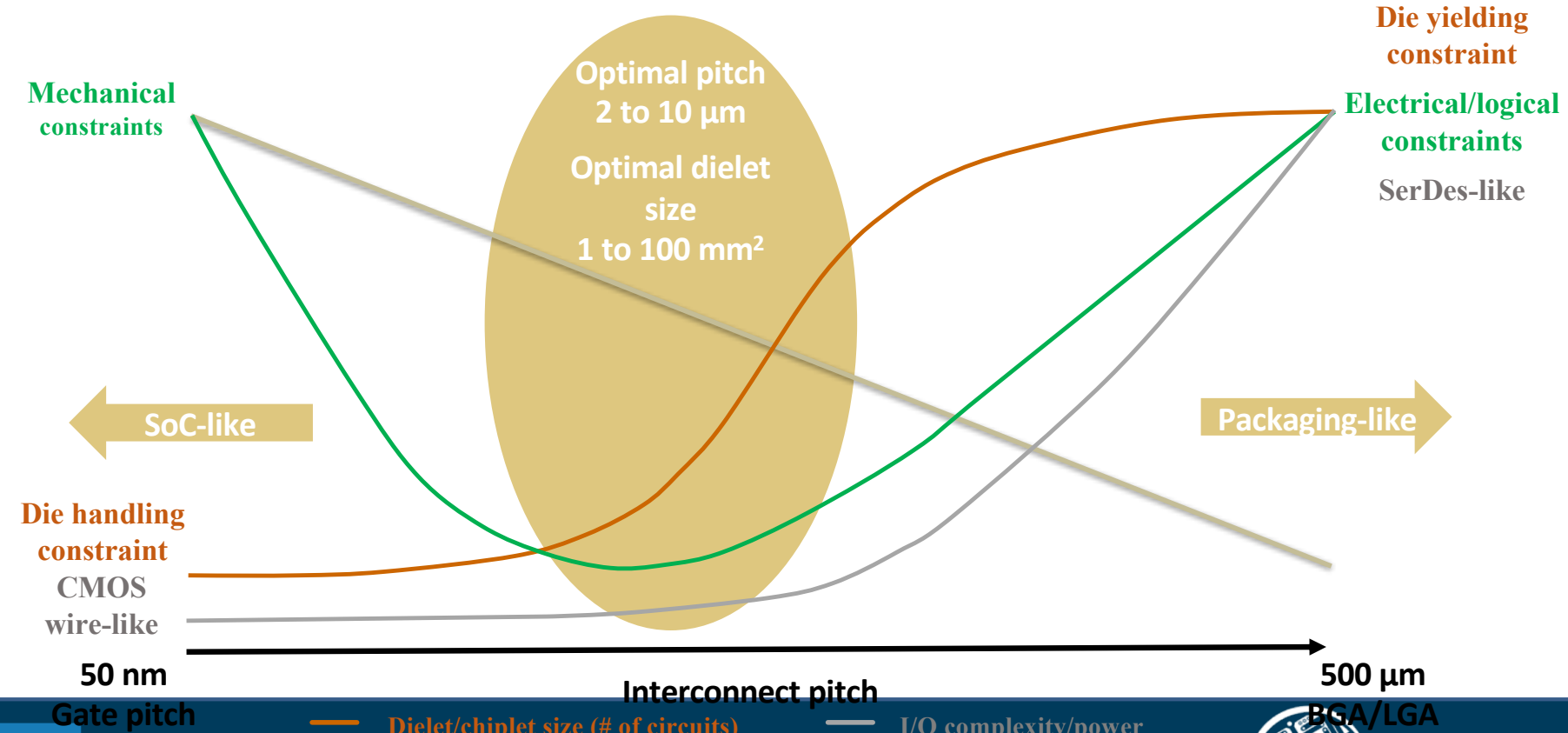
**We will use these terms interchangeably**

# Three Important Questions

- What is the optimal pitch at which dies should be interconnected ?
- What is the optimal dielet size
- How close should we assemble dies

Hint: how do we make a SOW look like an ginonormous SOC

# The Dielet Golden Regime

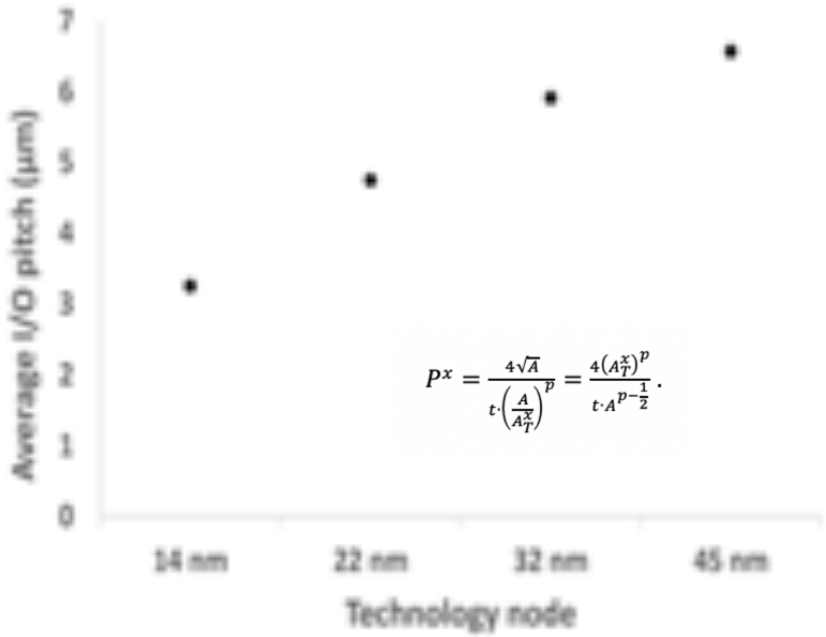



UCLA Samueli School of Engineering

CHIPS CENTER FOR HETEROGENEOUS INTEGRATION AND PERFORMANCE SCALING

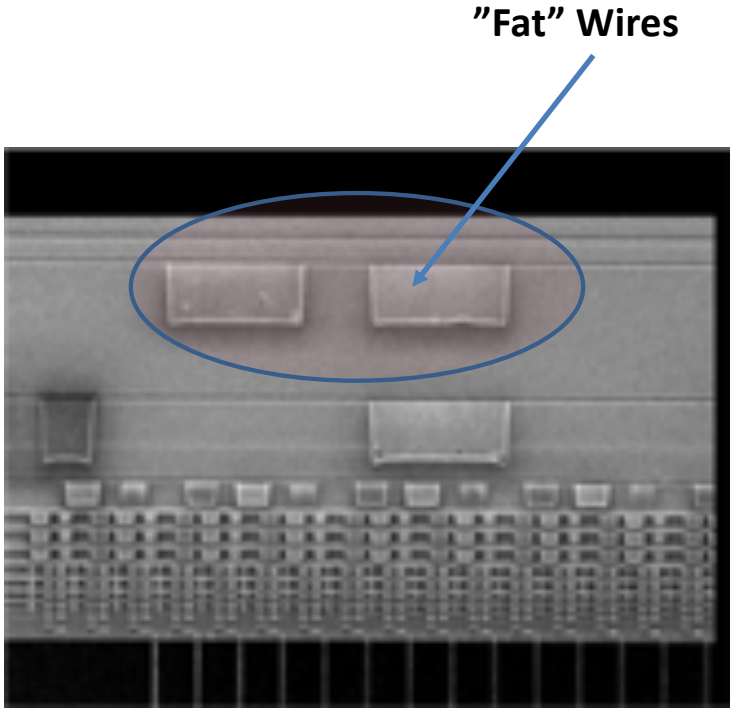
# What is the optimal I/O pitch ?

| Chip | Area ($mm^2$) | Transistor count ($x10^9$) | Technology node (nm) |
|---|---|---|---|
| IBM POWER9 [26] | 695 | 8 | 14 |
| AMD Zen [27] | 44 | 1.4 | |
| IBM POWER8 [28] | 649 | 4.2 | 22 |
| Intel Xeon Haswell E5 [29] | 663 | 5.56 | |
| IBM POWER7 + 80 MB [30] | 567 | 2.1 | 32 |
| Intel Itanium Poulson [31] | 544 | 3.1 | |
| IBM POWER7 + 32 MB [32] | 567 | 1.9 | 45 |
| Intel Xeon 7400 [33] | 503 | 1.9 | |



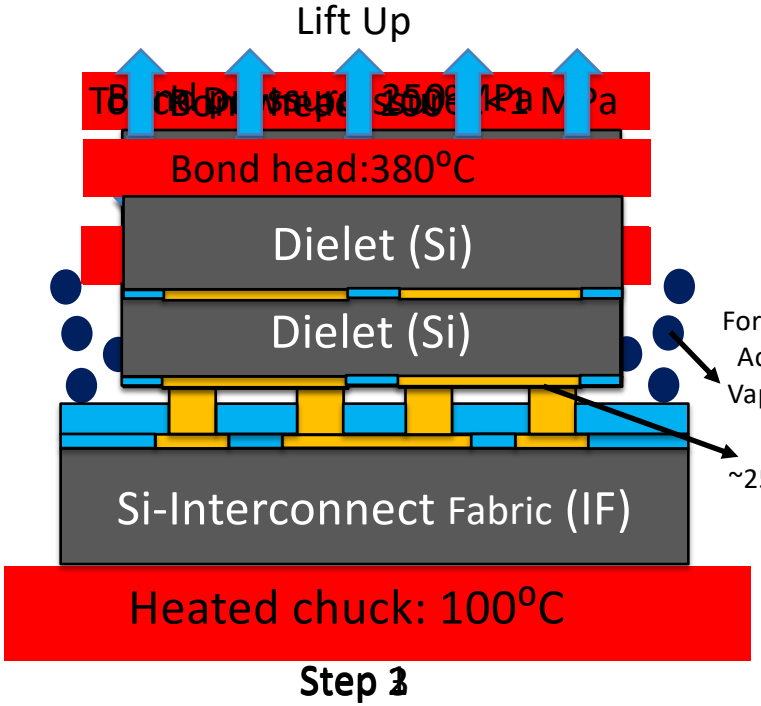
# Some Rationalization

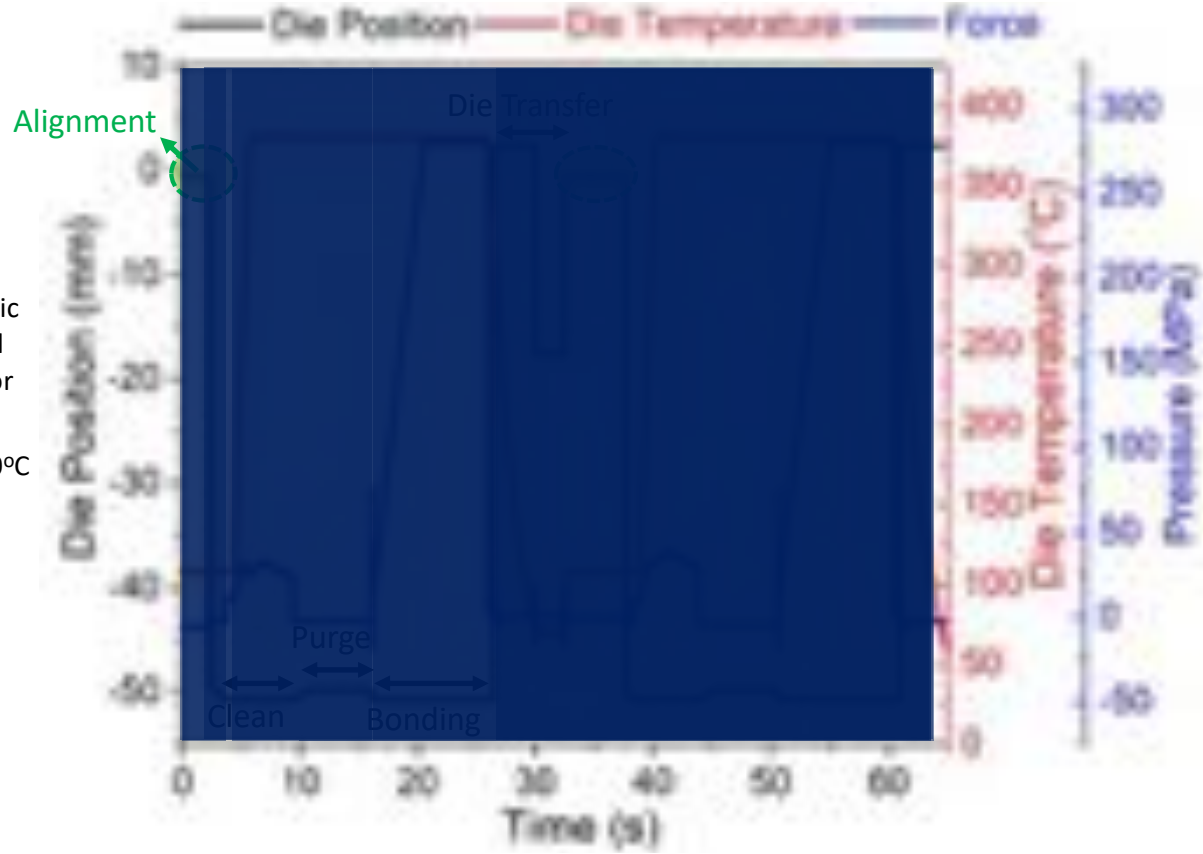
- Fine pitch ?
  - like ”fat wires” on a Silicon wafer - 2-10 μm
- Precision alignment ?
  - similar to fat wire alignment <0.2 μm
- Close Spacing
  - As close as possible <20 μm
- Typical block sizes on an SoC are typically a few ~100 μm on a side



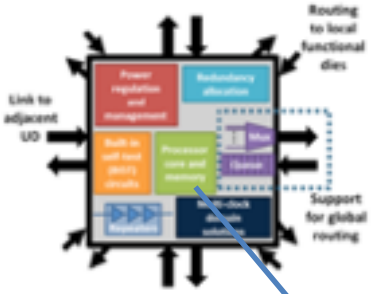

Hierarchical on-chip wiring system

# Direct Cu-Cu TCB process



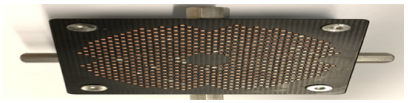



# System-on-Wafer (SoW) component Technologies

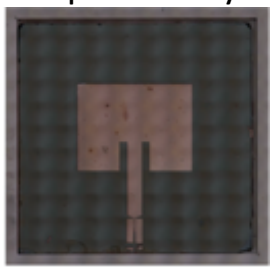


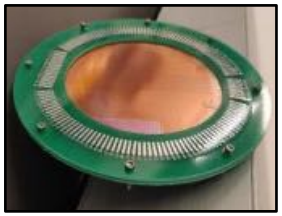

Network/utility die on IF

Single+/Two phase thermal solution

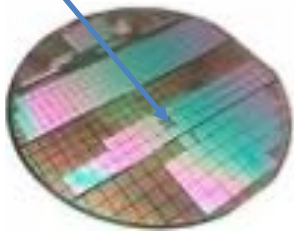


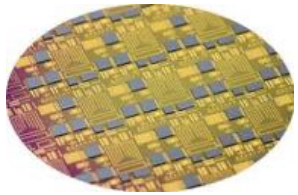
5G patch antenna on quartz inlay





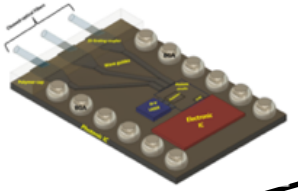
Connector collar

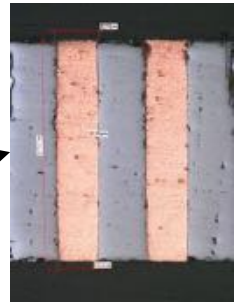




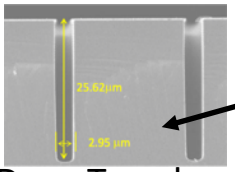
Populated Si-IF (Si (memory, Logic) and III-V dies)
With ALD AlN passivation

Photonic interconnect





Thru wafer Vias (TWVs



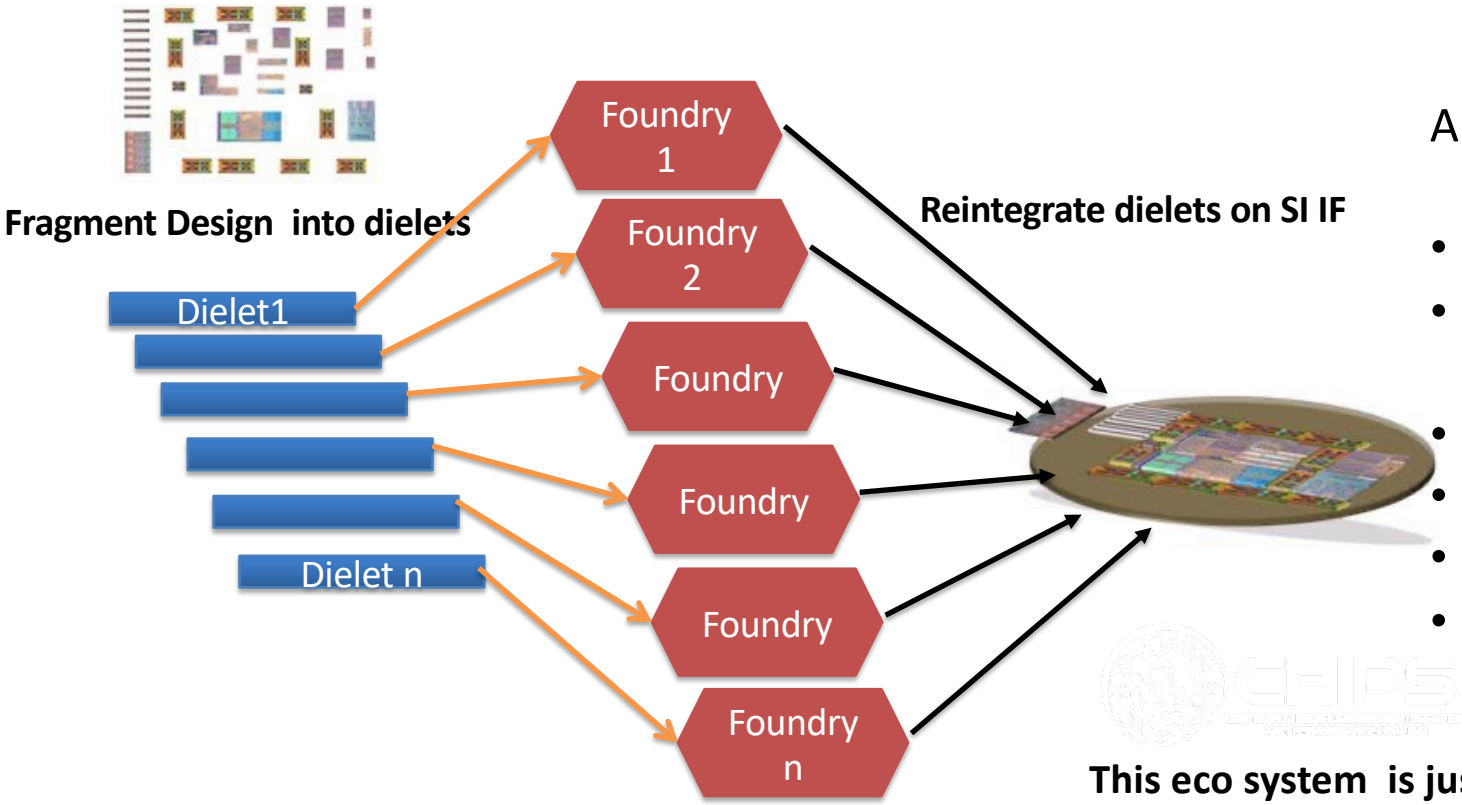

Deep Trench decoupling caps >250nF/mm²

Power board

Low Inductance CTE FlexTrate-based Coolable CTE compliant connector

$Z_{PDN} < 10m\Omega$

# Dielet fragmentation and reintegration scheme




Advantages:

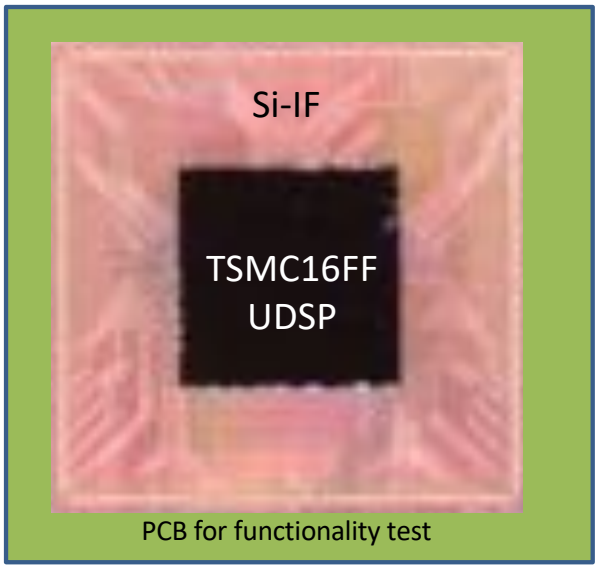
- Short Time to Market
- Low Non-Recurring Engineering (NRE)Costs
- IP reuse
- Extreme customization
- Extreme Heterogeneity
- Greater Yields die to small dielets

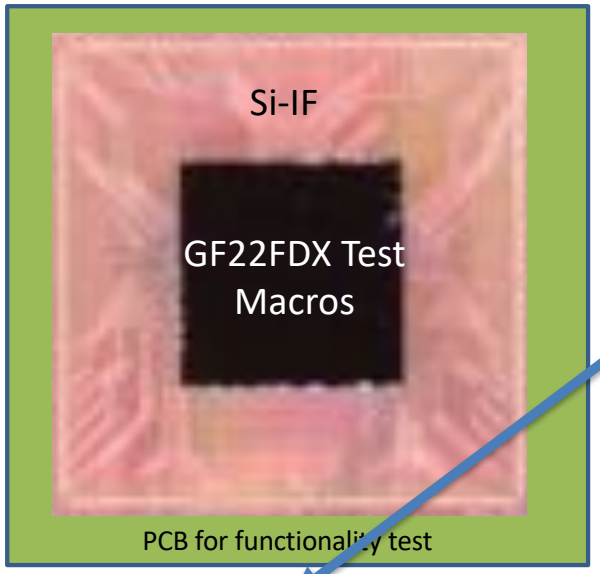
**This eco system is just beginning to take root**

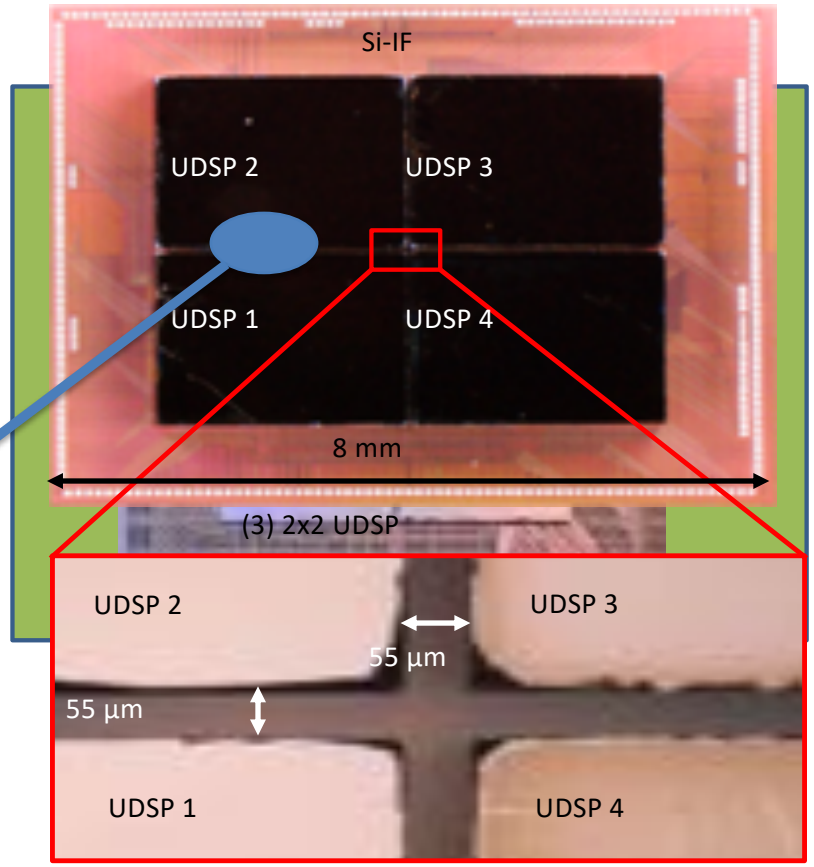
# Some Key requirements

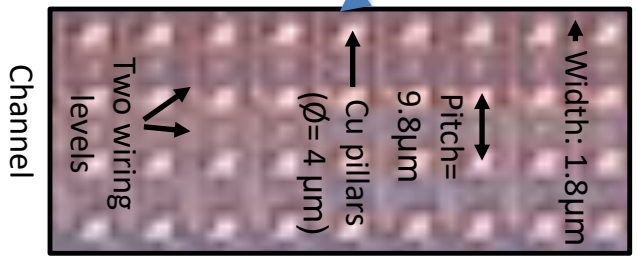
- Supply Chain Security  especially dielet availability ?
- Dielet standards
  - Mechanical : dimensions, terminations, thickness, edge preparation
  - Electrical: pitch, shorelines, electrical drivers, ESD, communication protocols
- Assembly houses
- Reliability

# Assembly on SiIF



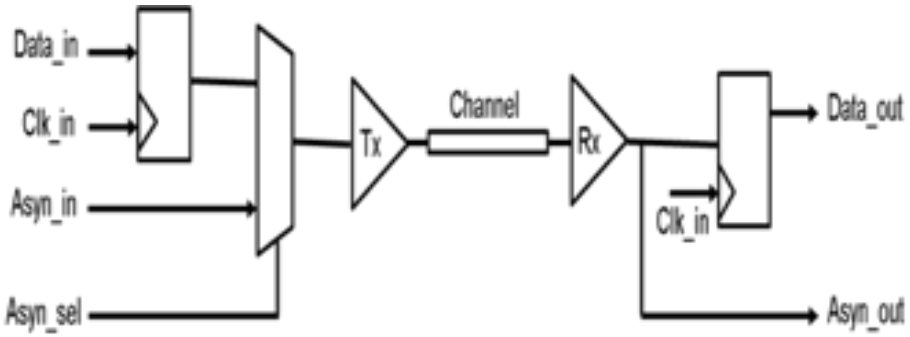




Measured Latencies:
21 ps for a 500 mm link
11 ps for a 200 mm link

# Inter die protocols become simpler as we parallelize more




Eg. SuperCHIPS protocol

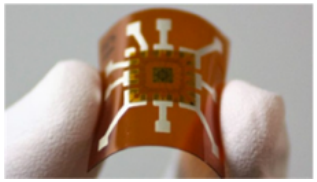
Design specifications

- Inter-DSP Bandwidth: 250 Gbps/mm
- Data-rate/link: 1 Gbps
- Latency: <70 ps/ 2 clock cycle
- Energy/bit:
  - 0.04 pJ/b (Async)
  - 0.2 pJ/b (sync)

# Dielets are in use in Flexible Hybrid Electronics today

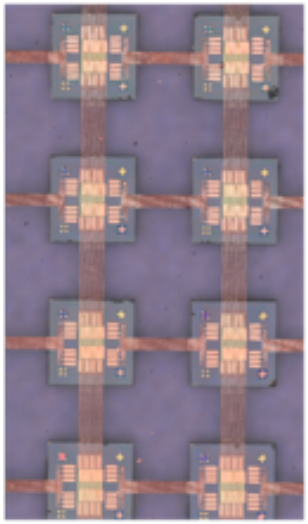
**Gen I FHE Today**


*Source: Uniqarta*



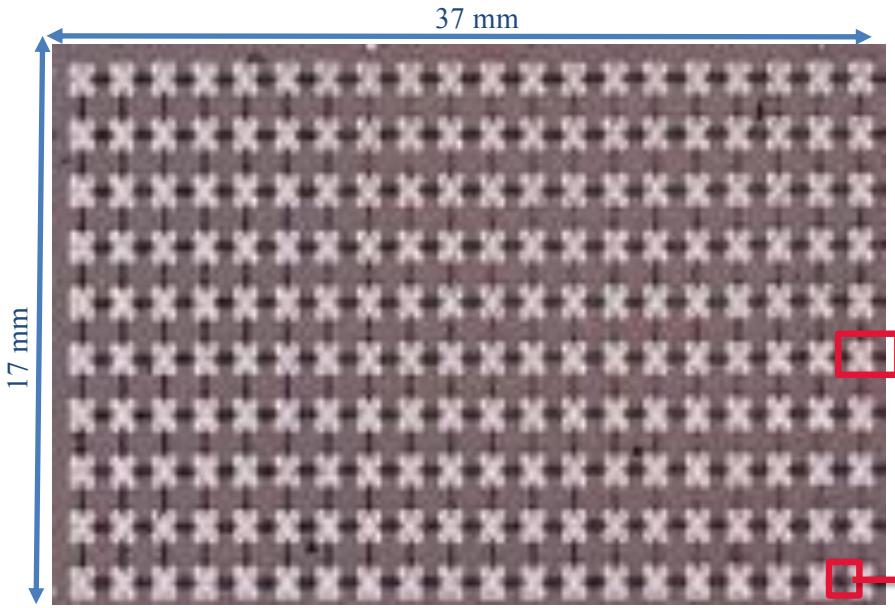
Printed electronics & wiring, ultra-thin die, organic thin film semiconductors - at best bendable

**Gen II FlexTrate™**

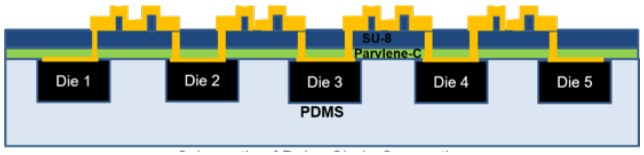


- Extremely flexible, even foldable and stretchable
- FOWLP (Fan-Out Wafer-Level Packaging)approach
- Lithographically defined fine pitch corrugated interconnects
- Allows for hetergogeneous dielet integration including Thin-film batteries, passtives. connectors

# Dielet approach to flex



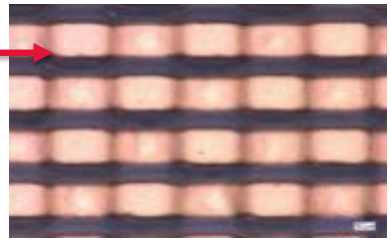

10×20 Array of Dies




Schematic of Daisy chain connection


W/S: 20µm/20µm

Die



W/S: 27µm/13µm

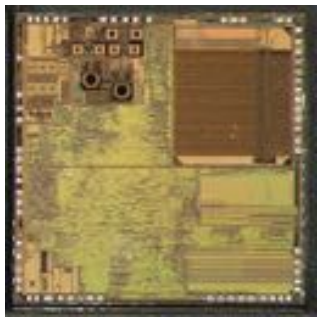
Vertically corrugated interconnects

- **1 mm$^2$ dies**
- **1.8 mm die pitch**
- **Vertically corrugated interconnects allow fine pitch integration**

# Bluetooth Size Reduction


7 mm × 7 mm × 0.85 mm

**QFN Package to Bare Die**



3.5 mm × 3.5 mm × 0.2 mm

- **4x** area reduction
- **17x** volume reduction
- Many companies such as **i3 Microsystems** and **American Semiconductor** provide chip decapsulation

# Summary

- Heterogeneous integration is not really new…….
- But the paradigm is changing
  - Dielets need to be small
  - Dielets need to be connected at fine pitch
  - Dielets need to be assembled very close to each other
  - As pitch becomes smaller paralyzed protocols are easier to standardize
- The flex/medical device industry is embracing the dielet concept but with somewhat *ad hoc* solutions

- Several challenges remain but are not really showstoppers
- We need to develop an ecosystem for dielets and a secure supply chain

# UCLA CHIPS Thanks you for your Support !


UCLA Samueli School of Engineering

CHIPS
CENTER FOR HETEROGENEOUS INTEGRATION AND PERFORMANCE SCALING

# Acknowledgements

https://chips.ucla.edu/students

Students
Arsalan Alam
Pranav Ambhore
Arpan Dasgupta
Goutham Ezhilarasu
Randall Irwin
Siva Jangam
K.T.Kannan
Guanxi Oyang
Saptaeeep Pal
Gholam Sabbir
Krutikesh Sahoo
Ujash Shah
Niloo Shakoorzadeh
Frank Wan
Yu Tao Yang
Michael Molter
Fred Chu

Postdocs:
Samatha Benedict
Umesha Mogera

Alums:
Adeel Bajwa
Amir Hanna
Boris Vaisband
Tak Fukushima

Andy Liu
Eric Sorenson
Arshiya Vohra
Mohd Abbas