Designing Scale-out Chiplet Based Systems

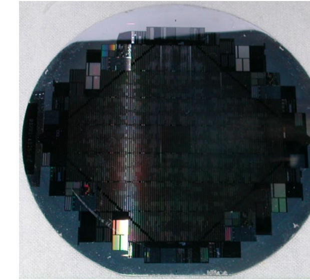
PUNEET GUPTA

ECE DEPT., UCLA

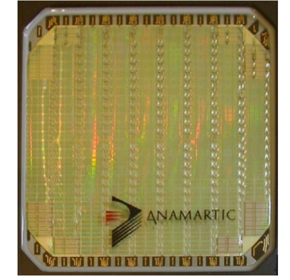
PUNEETG@UCLA.EDU

A Brief History of Waferscale Computing

- Waferscale systems: processors that span a full silicon wafer
  - 100mm wafer ~ 7900mm2
  - 200mm wafer ~ 31,400mm2
  - 300mm wafer ~ 70,000mm2
  - Comparison: largest System on Chip ~ 800mm2



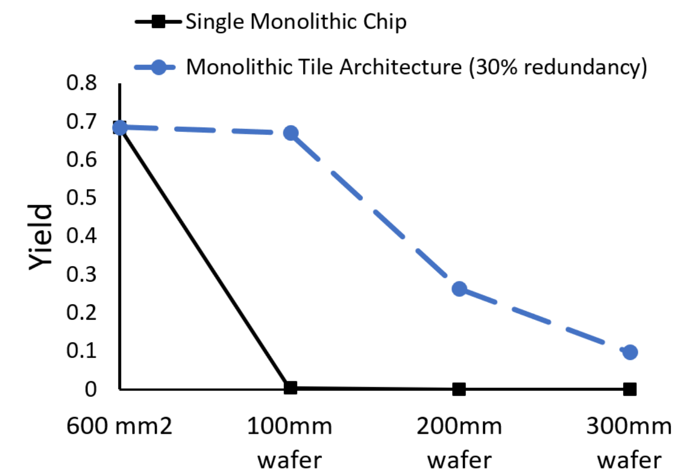
Trilogy Systems



Tandem Computers, Fujitsu

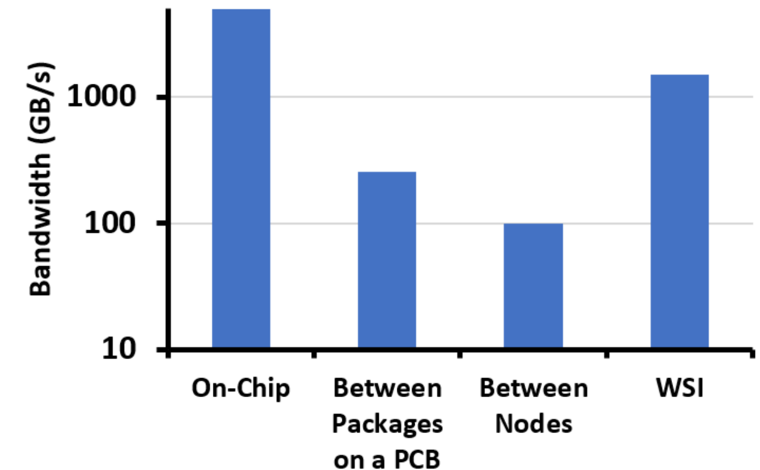
Other efforts: ITT Corporation, Texas Instruments. Recent efforts: Spinnaker (Neuromorphic Chip), Cerebras, Tesla

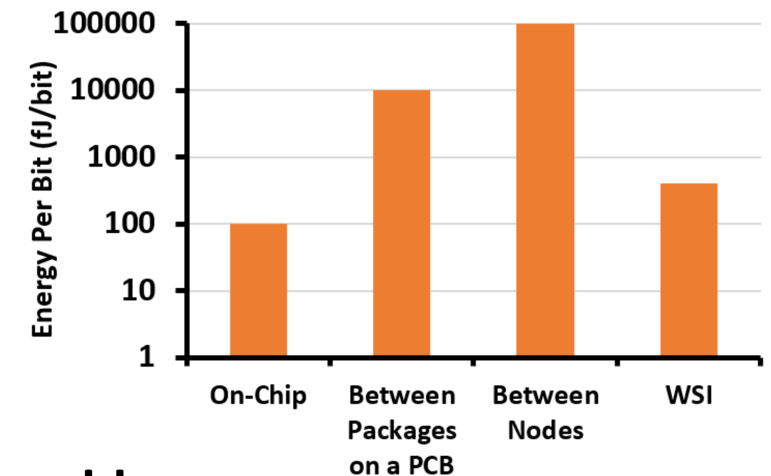
- Waferscale processing has been discussed since 1980s
  - Never really worked out due to yield issues
  - Recent efforts more promising



Time to Give Waferscale Another Go?

- Highly parallel applications are spread across many processors
- Communication between the processors is still a big bottleneck
  - Low Bandwidth (a few 100s of GBps)
  - High energy per bit (10s of pJ/bit)
  - Real estate on chip (15-25% of the chip is devoted to SERDES I/Os)



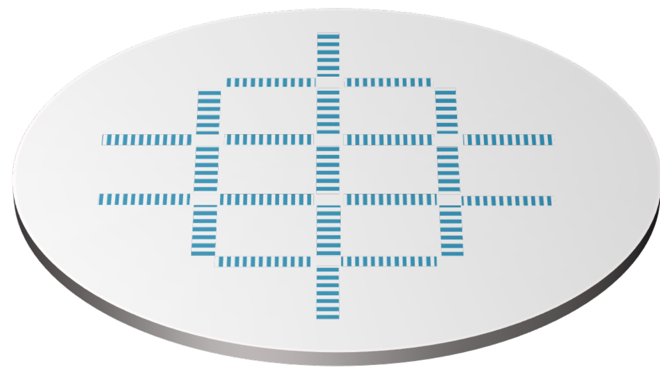


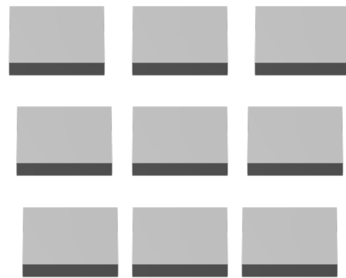
However, to achieve scale-out integration, we need to solve the **yield problem**

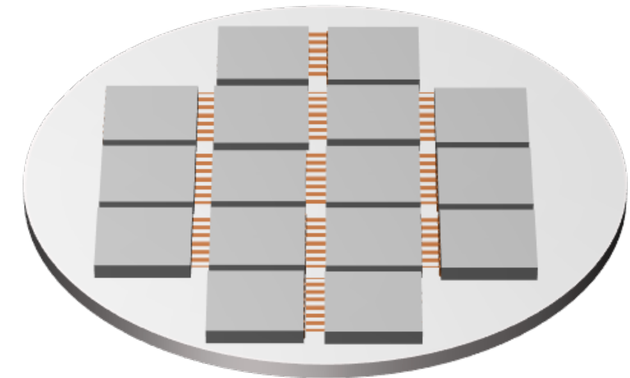
Re-imagining Waferscale Integration

Q: What do we need from waferscale integration?

A: High density interconnection







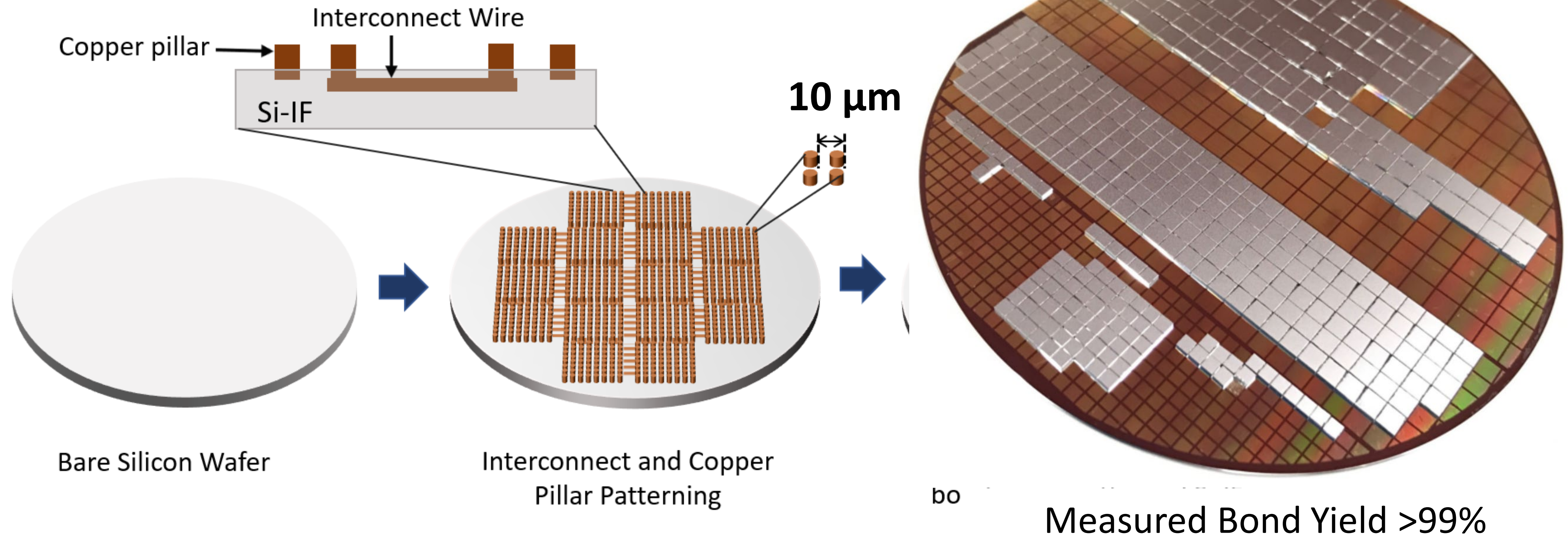
A wafer with interconnect wiring only

Small known good dies

Bond the dies on to the interconnect wafer

Enabling WSI Technology

UCLA Silicon Interconnect Fabric (Si-IF)*



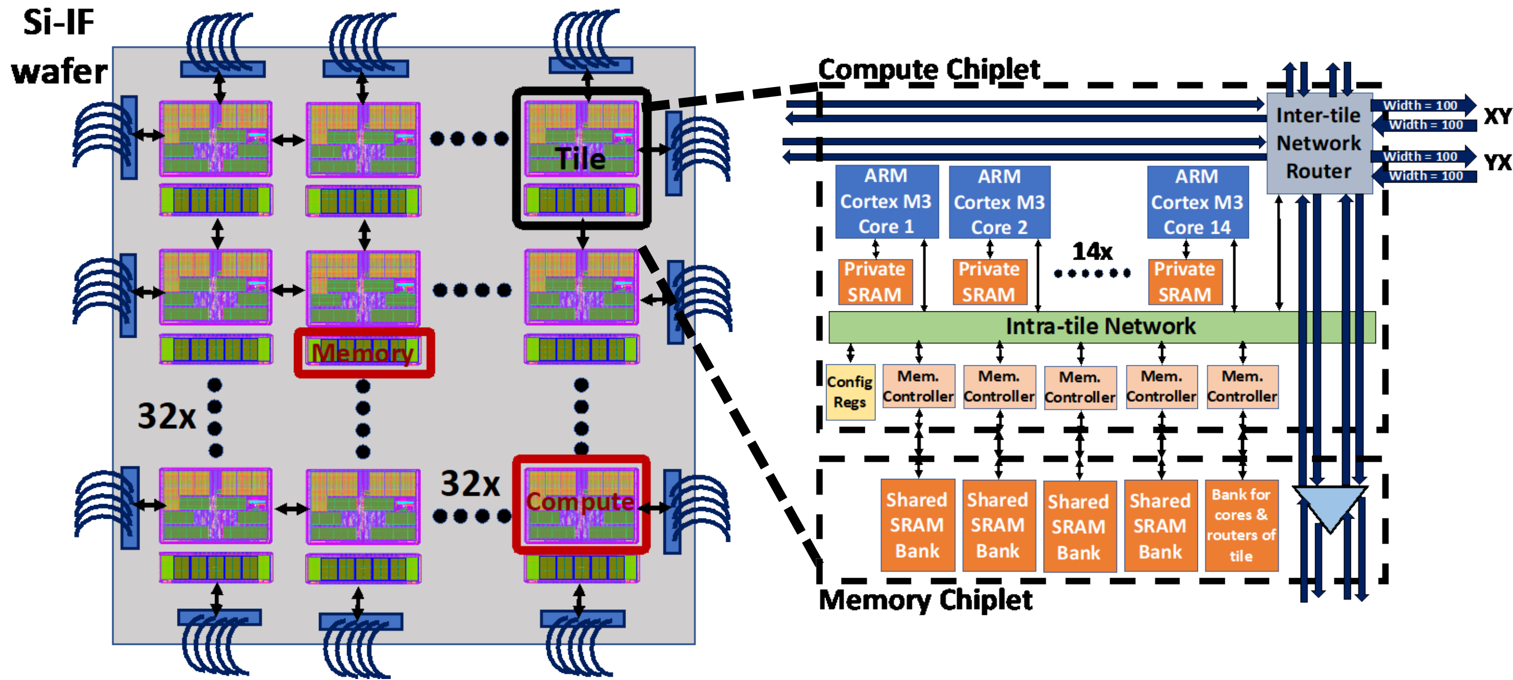
Allows waferscale integration with high yield

*UCLA CHIPS Program: https://www.chips.ucla.edu/research/project/4

Designing a Waferscale Processor Prototype: Challenges and Solutions

[Appeared in DAC'21, ECTC'21]

2048 Chiplet Architecture



- Implemented in **TSMC N40-LP**
- Tiles : **1024 (Total 14,336 Cores)**
- Private memory per Core: **64KB**
- Total Shared Memory: **512MB**
- Shared Memory Bandwidth : **6.14 TB/s**
- Network Bandwidth: **9.83 TB/s**
- Total Compute : **4.3 TOPs**
- Peak Power : **725W**
- Total Area: **15100 mm²**

Challenges Faced While Designing the System

How should we **deliver power** to all the flip-chip bonded chiplets across the wafer?

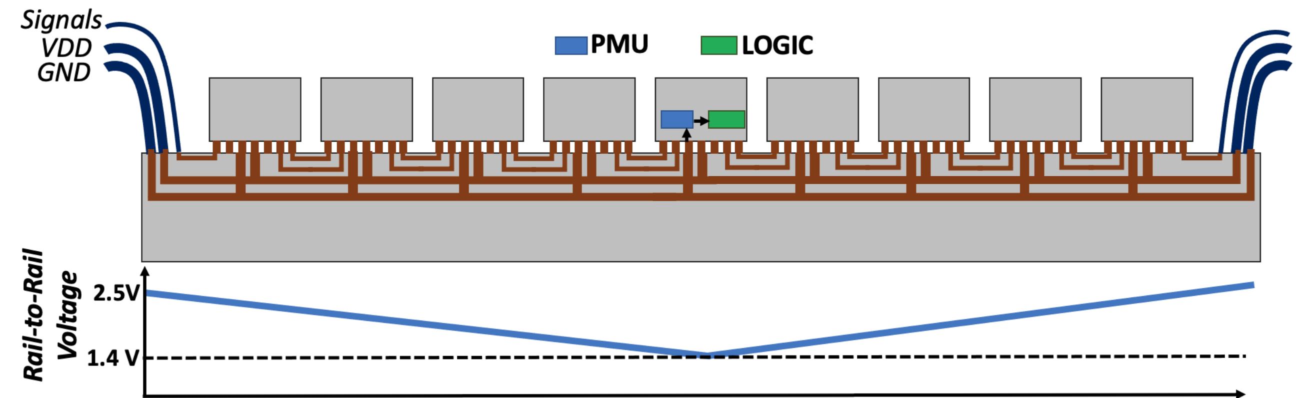
How can we reliably **distribute clock** across such a large area?

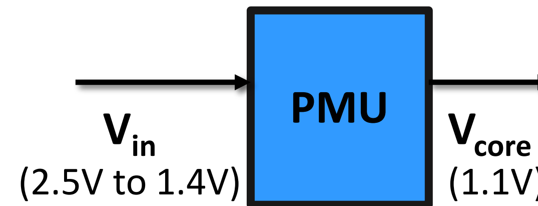
What is the **testing strategy** for such a large system?

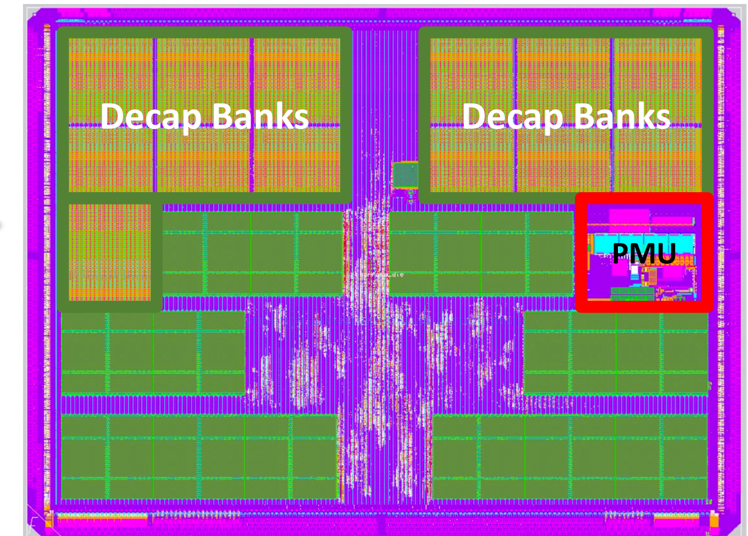
What is the **inter-chip network architecture** and how do we achieve resiliency if a few chiplets fail?

Power Delivery

- Edge Power Delivery at 2.5V
  - Wasteful but simple
  - Back or front side power delivery possible but more complex
- LDO based power management at each node
- On-chip decoupling capacitance (20nF per tile)
- DeCap consumes 30% of the chip area
  - *Deep Trench Capacitors would help*

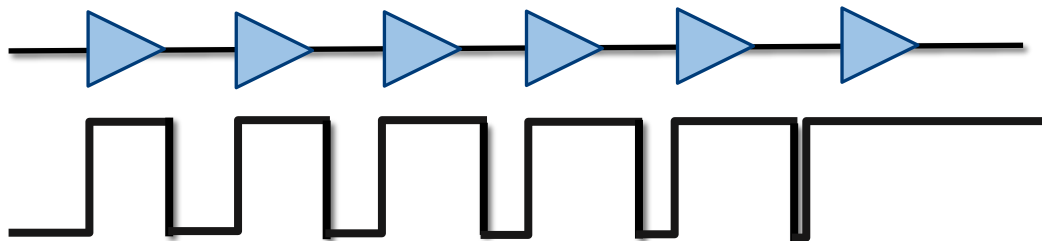




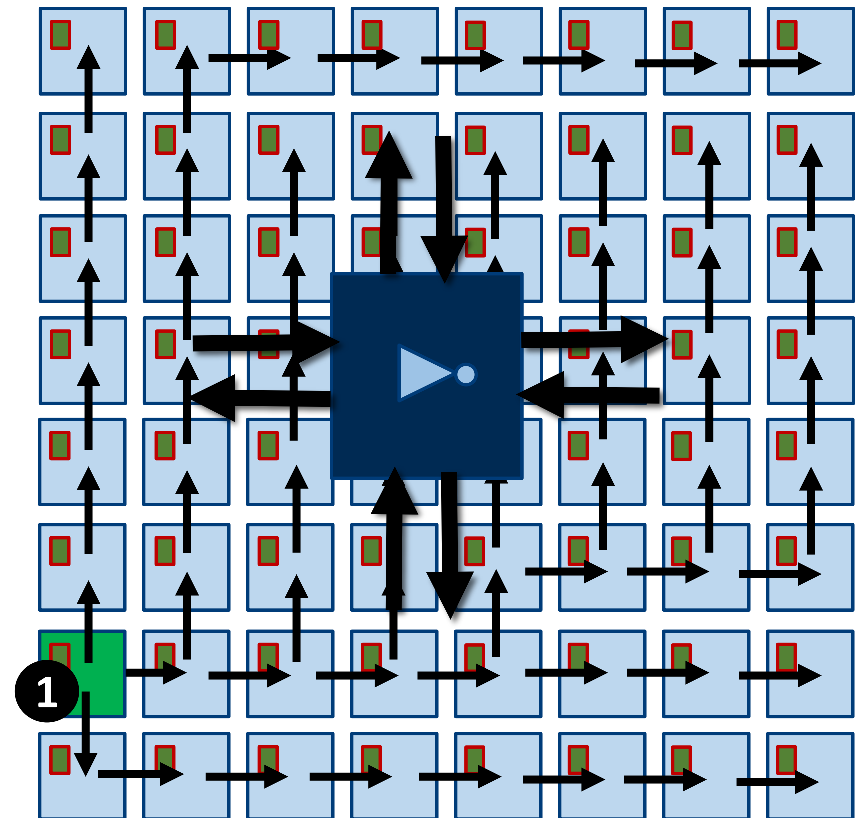


Waferscale Clocking

- Clock generation
  - Stable reference voltage needed by PLL not present away from edge
  - Generate fast clock at the edge and distribute
- Clock distribution
  - Fast clock is forwarded
  - Clock inverted at each hop to avoid duty cycle distortion accumulation

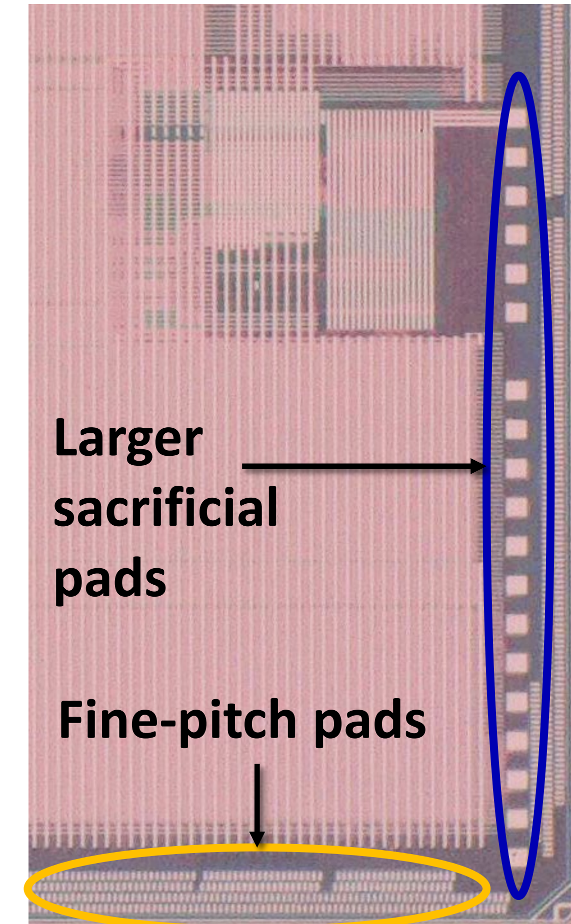


  - Communication between dies using asynchronous interfaces
  - Fault tolerance in clock distribution network



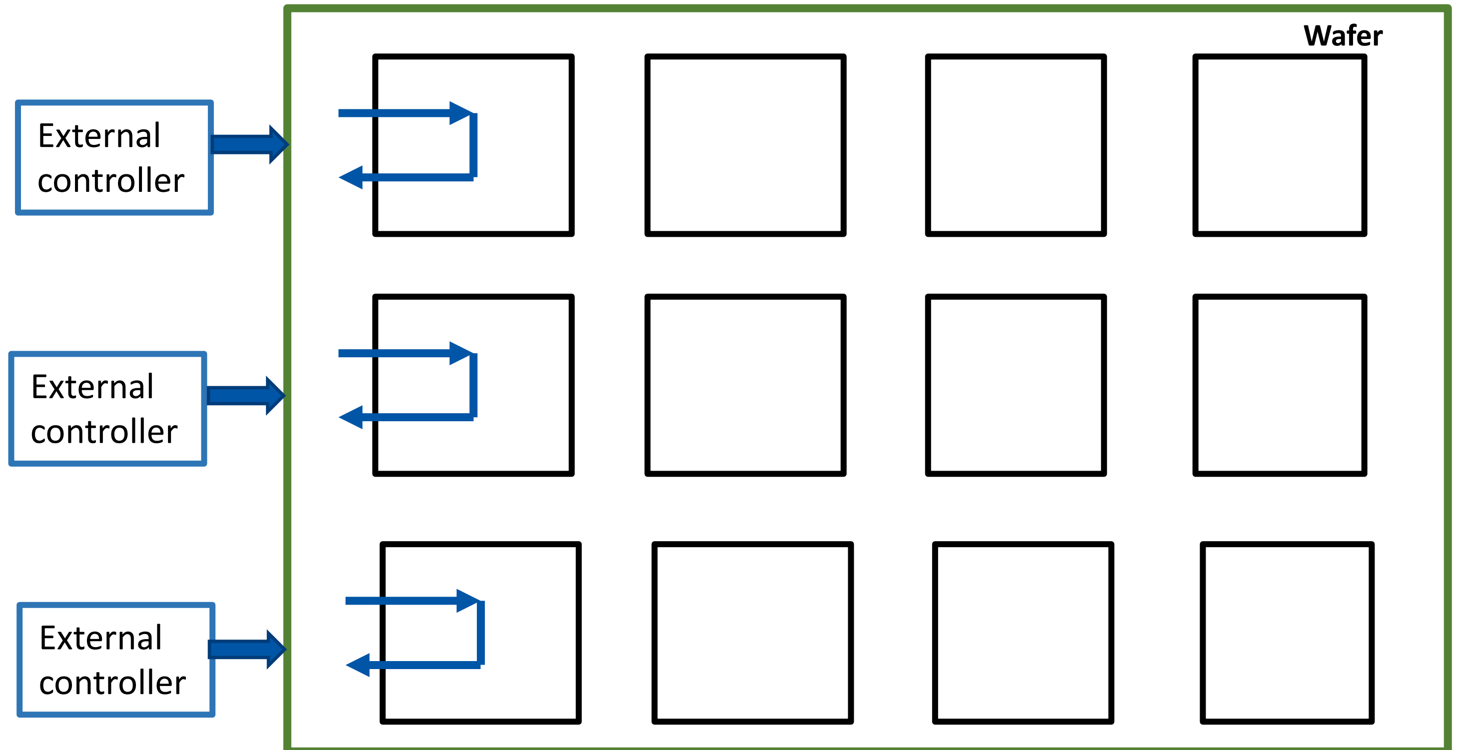
Pre-bond Die Testing

- Fine pitch pads cannot be probed
- Larger sacrificial pads for probe test

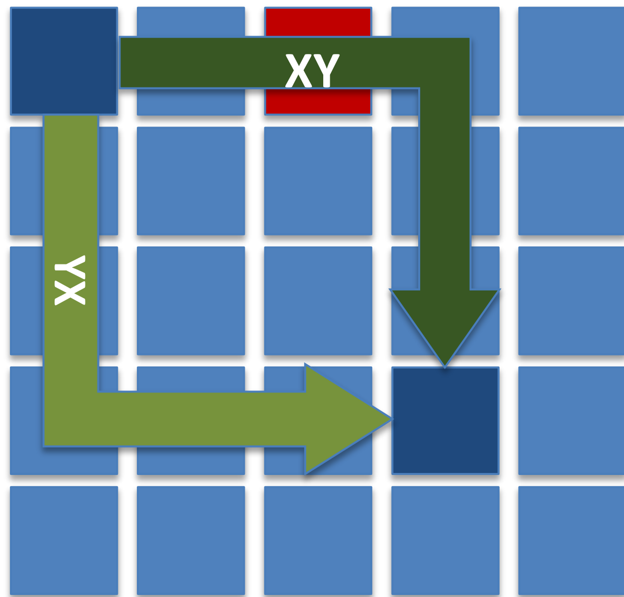


Post-bonding JTAG Test Scheme

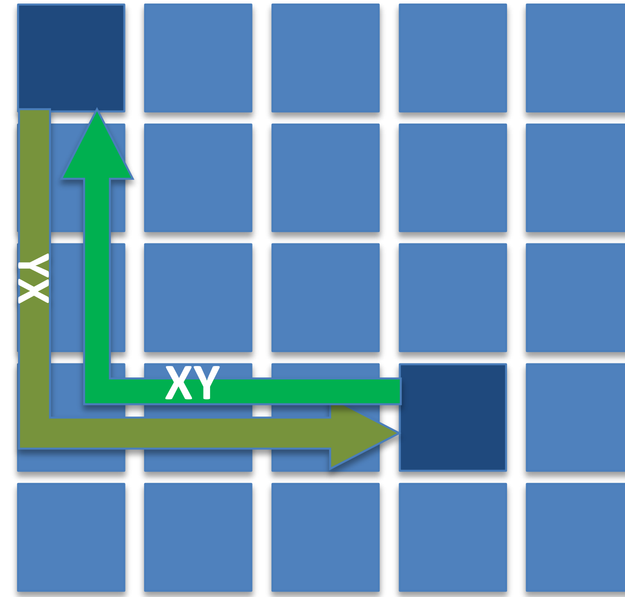
- (1) Multiple chains
  - One JTAG chain results in single point of failure vulnerability
  - Throughput is an issue:
    - 2.5 hours to load the memories using one chain
    - 5 minutes to load with 32 chains
- (2) Progressive unrolling
  - Helps identify post-bonding faulty dies
  - Similar to IEEE 1838 proposal



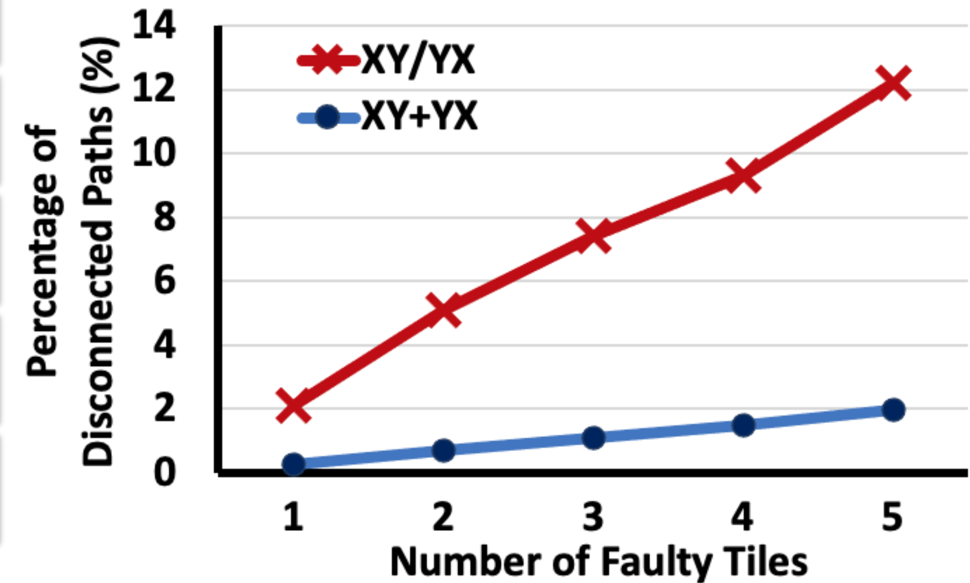
Network Resiliency



Two Separate Networks

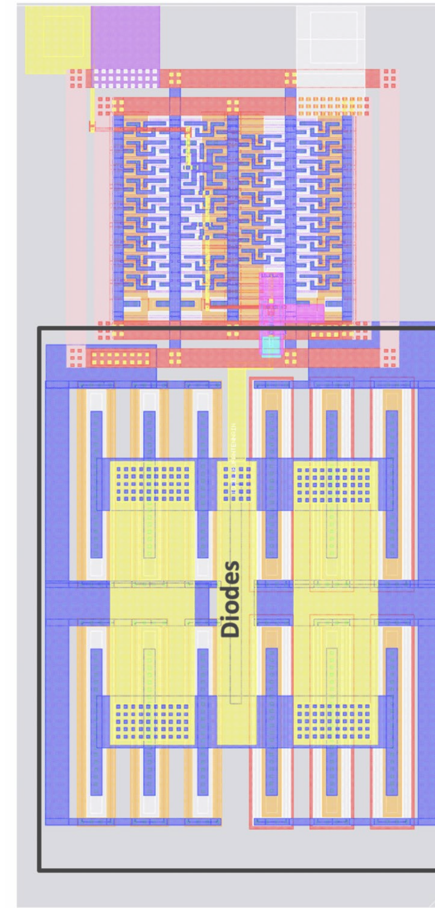


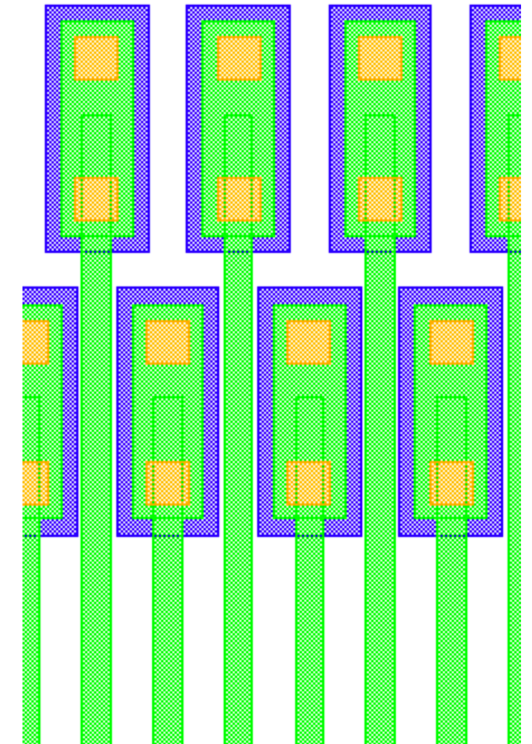
Request-Response in Complimentary Networks



I/O Architecture

- I/O pitch of 10 um and depth of 20 um
- Simple cascaded buffer architecture
- 0.07 - 0.18 pJ/bit
- Two pillars per IO for redundancy
- ESD diodes and buffers need to fit within the I/O footprint





Chiplets Tested, Waferscale System Assembly in Progress

3.18 mm

3.6 mm

Tile

3.18 mm

2.47 mm

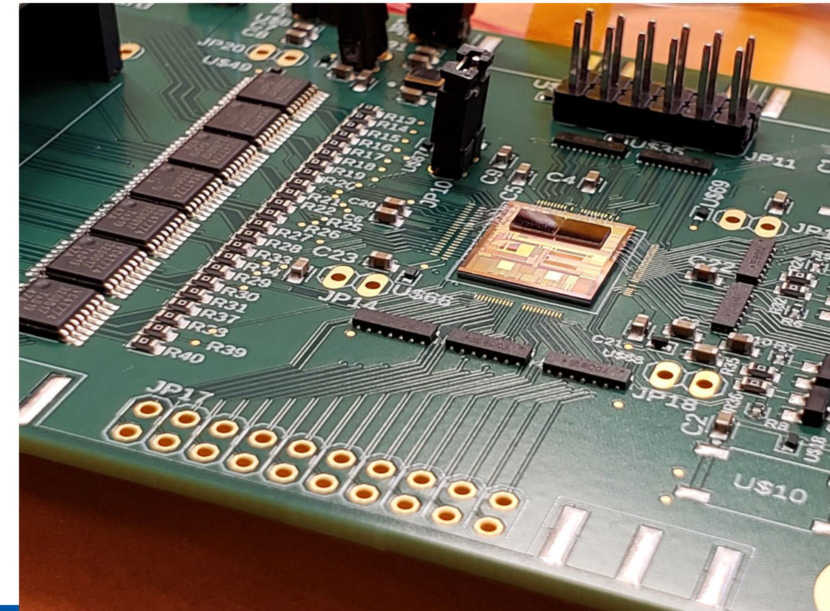
Compute

Memory

1.12 mm

3.18 mm

100μm



Summary

- Large scale chiplet-based processing is here to stay: chiplet advantages at the system-level

1. *Heterogeneity.*
   - Integrate logic non-compatible memories + network interfaces
   - Selective upgrades to system IPs
2. *Scale*
   - Ability to build large systems without yield concerns
   - Possibility to go beyond reticle size boundaries (e.g., waferscale)

- Scale-out waferscale chiplet-based design requires new (co)design methodologies: new ways to look at (old) problems
  - Test, reliability and fault tolerance
  - Physical design
  - Power distribution
  - Heat extraction

Acknowledgements

- CDEN, CHIPS centers and Qualcomm Innovation Fellowship for funding
- Students: **Saptadeep Pal**, Irina Alam, Krutikesh Sahoo, Shi Bu, Haris Suhail, Nick Cebry, Jingyang Liu, Daniel Petrisko, Matthew Tomei
- Collaborators: UCLA: Subramanian S. Iyer, Sudhakar Pamarti. UIUC: Rakesh Kumar