zerlol

Building Composable Chiplets

CHISIC Design Workshop
Apr 16th -17th, 2025
Boston MA

Andreas Olofsson
andreas@zeroasic.com

Chip(let) lessons from the trenches

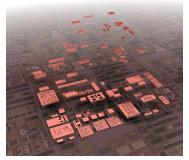


Test:
1. ATEs are expensive and slow moving walled garden.
2. BIST is the future

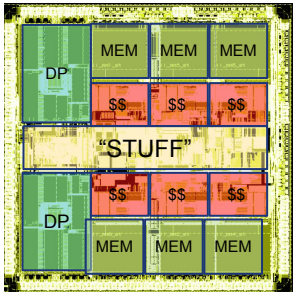
ASIC[2]:
1. ASICs are easy
2. ASICs are fun
3. Agile is the future



DARPA R&D[5]
1. Chiplets are the future
2. Aim high
3. 10 year R&D cycles
4. 9 out of 10 efforts fail



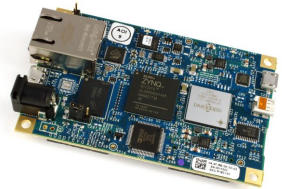
1996 — 1998 — 2006 — 2008 — 2017 — 2020

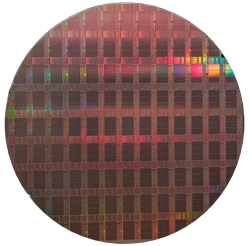


SoC[1]:
1. Complexity begets complexity
2. Large teams are inefficient
3. VLIW/eDRAM are bad ideas
5. 3% of SoC does useful work

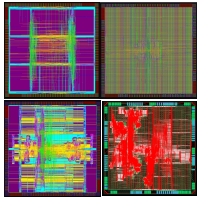


Manycore[3,4]:
1. Parallelism is the future
2. Networks (NoCs)
3. Automation
4. Openness





Chiplets:
1. Semiconductor is broken
2. Composability is the future
3. Digital twins are the future
4. Open source is the future

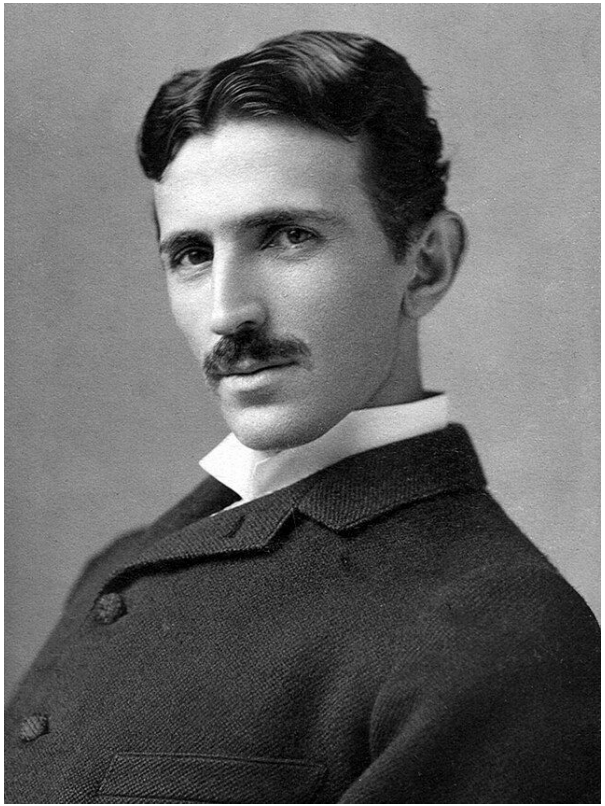


[1] Adelman, Olofsson et al (2004), "A 600 MHz DSP with 24 Mb embedded DRAM with an enhanced instruction set for wireless communication", ISSCC
[2] Olofsson et al (2008). A variable width software programmable data pattern generator (U.S. Patent No. 8,006,114)
[3] Olofsson et al (2011) A 1024-core 70 GFLOP/W floating point manycore microprocessor, High Performance Embedded Computing Conference
[4] Olofsson et al (2014), "Kickstarting high-performance energy-efficient manycore architectures with Epiphany" 48th Asilomar Conference on Signals, Systems
[5] Olofsson et al (2018), Enabling High-Performance Heterogeneous Integration via Interface Standards, IP Reuse, and Modular Design, IMAPS

Startup Lesson #1

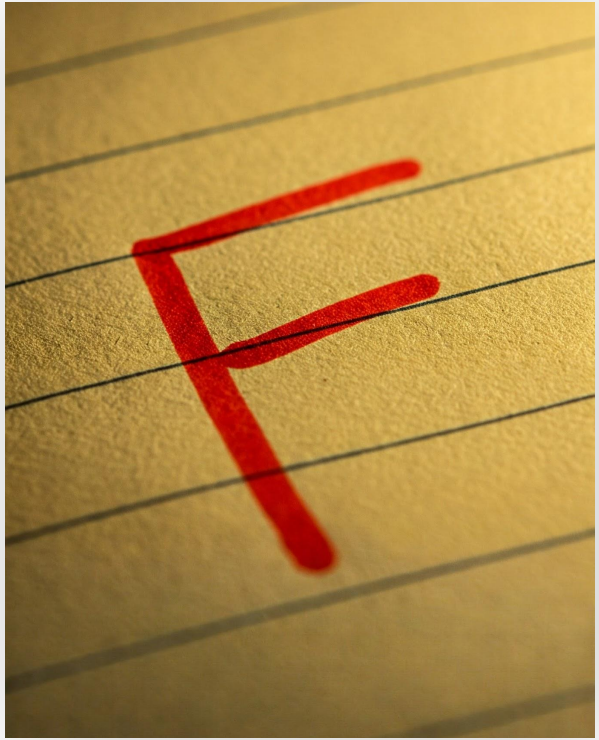
“Den som är före sin tid får invänta framtiden på en obekväm plats.” – Lennart Lubeck, CEO Swedish Space Corporation (1980’s)

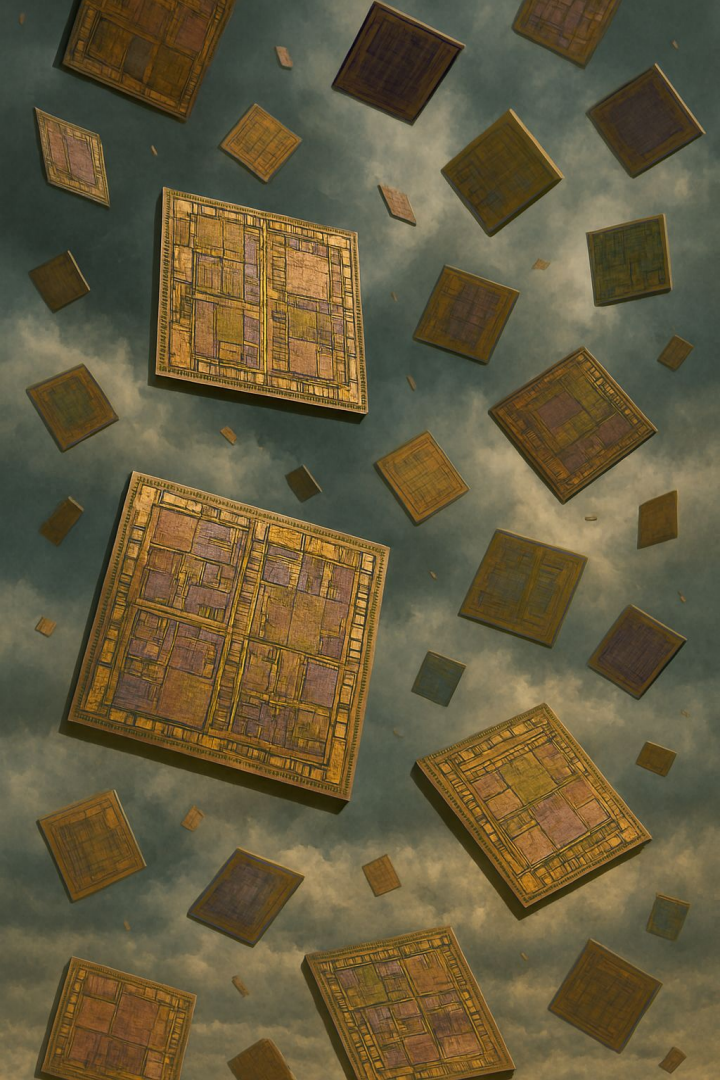
“Those who are ahead of their time often have to wait for it in uncomfortable quarters.” – Stanislaw Lec (1909-1966), Polish aphorist, poet



2025 Chiplet Report Card

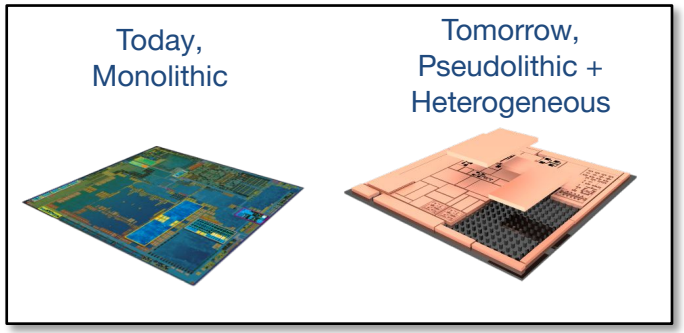
- **No** open access chiplets for sale
- **No** full-stack chiplet standard
- **No** solution to the PPM/KGD problem
- **No** solution to margin stacking problem
- **No** solution to the chip rework problem
- **No** 3rd party SOTA chiplet integrators
- **No** low volume advanced packaging
- **No** funding for chiplet ecosystem development
- **No** viable PCB like design ecosystem



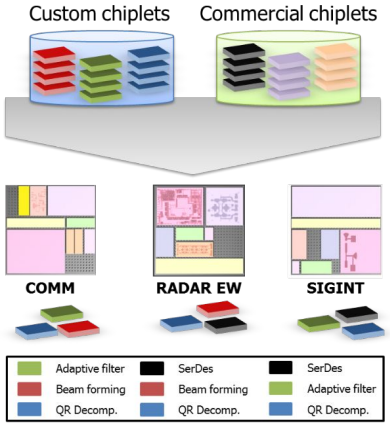


How Did we Get Here?

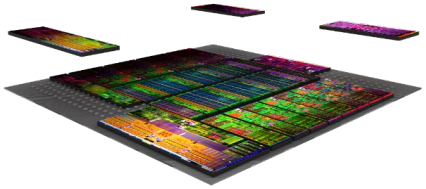
(2016) DARPA CHIPS Program Launch



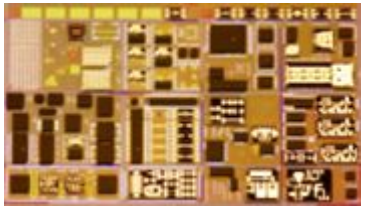
REF: DARPA @ Semicom Design West 2019



- ✔ A universal efficient interface standard
- ✔ SOTA manufacturing assembly
- ✔ A large and critical set of IP chiplets



Extend Moore's law
Scale out and scale down while managing yield

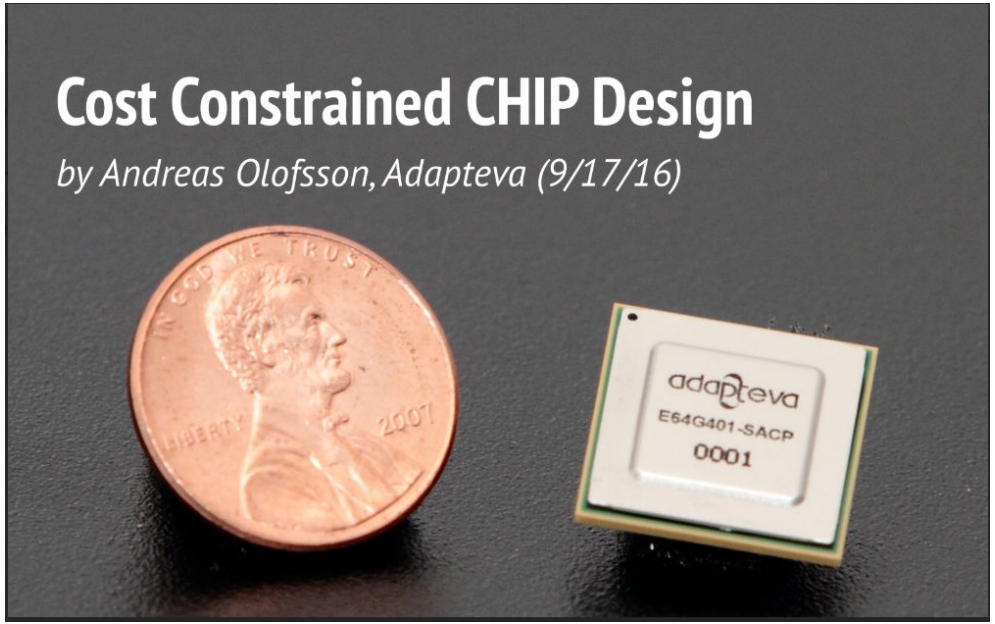


Heterogeneous Integration
Materials/processes, companies, geography, security



System Integration
Democratize access to leading edge silicon for system integrators

(2016) CHIPS Proposers Day



6

Our I/O Challenges

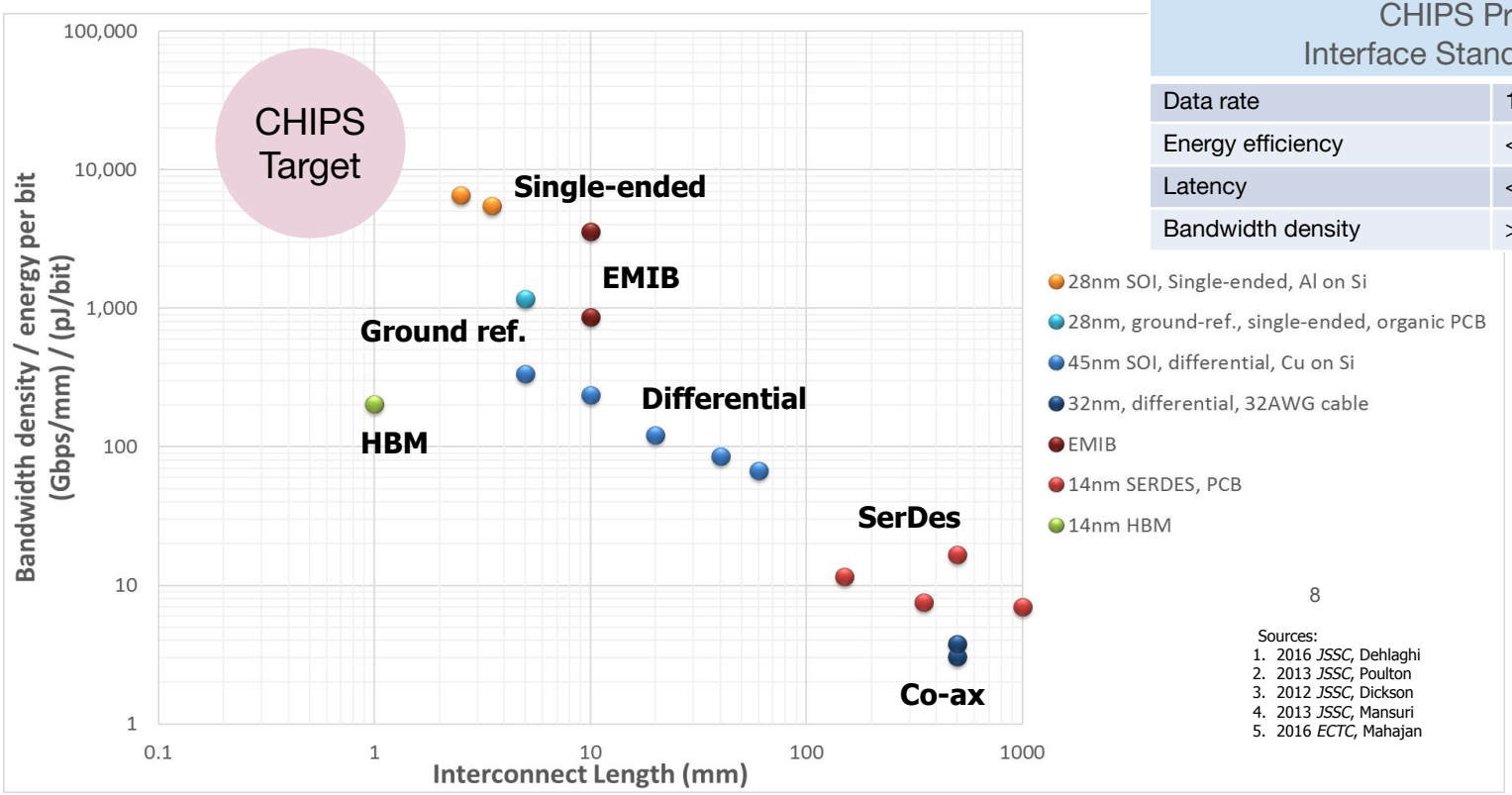
- Needed massive IO to support 4 TFLOPS!
- Initial plan was 128 x 10Gbit SERDES lanes
- ...but cost made monolithic integration impossible
- fallback was 1024 1.8V CMOS pins running at @ 150MHz...

8

One Possible Chip to Chip Interface

- Make sure 50um bumps are available to all customers
- Drive parallel interfaces (clk, frame, wait, data[N-1:0])
- Reference RTL: github.com/parallella/oh
- CMOS signaling using thin oxide transistors (0.8V)
- Energy Target: 0.2pJ / bit
- Density Target: 2Tbit / mm^2

(2017) CHIPS Standards War



CHIPS Program Interface Standard Metrics	
Data rate	10 Gpbs
Energy efficiency	< 1 pJ/bit
Latency	< 5 ns
Bandwidth density	> 1000 Gbps/mm

8

Sources:
1. 2016 *JSSC*, Dehlaghi
2. 2013 *JSSC*, Poulton
3. 2012 *JSSC*, Dickson
4. 2013 *JSSC*, Mansuri
5. 2016 *ECTC*, Mahajan

[REF] Olofsson et al (2018), Enabling High-Performance Heterogeneous Integration via Interface Standards, IP Reuse, and Modular Design, IMAPS

(2018) CHIPS Open source AIB FTW!

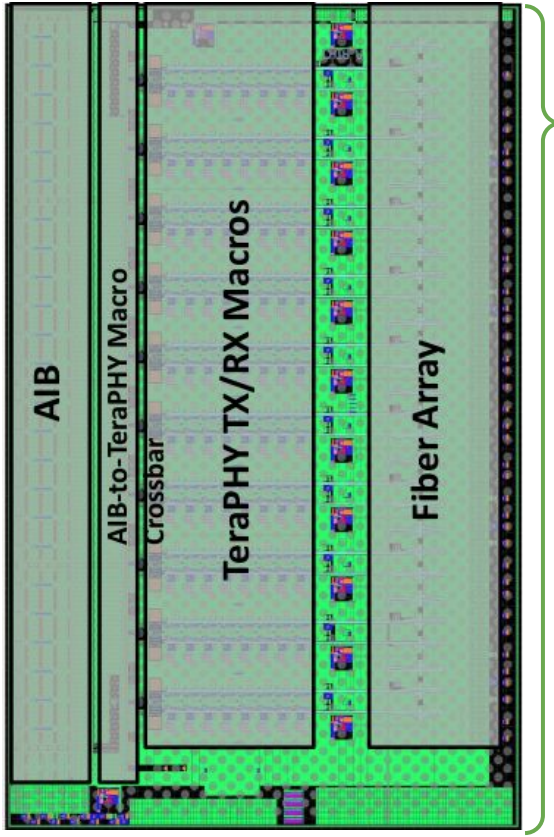
- **AIB (Advanced Interface Bus)** is a PHY-level interface standard for high bandwidth, low power die-to-die communication
  - Clock-forwarded parallel data transfer like DDR DRAM
  - High density with 2.5D interposer (e.g., CoWoS, EMIB) for multi-chip packaging
  - PHY level only (OSI Layer 1)
  - Protocols like AXI-4 can be built on top of AIB
- **AIB Performance:**
  - 1 Tbps/mm shoreline
  - ~1pJ/bit
  - <5ns latency
- **Open Source!**
  - Standard and reference implementation
  - https://github.com/chipsalliance/aib-phy-hardware

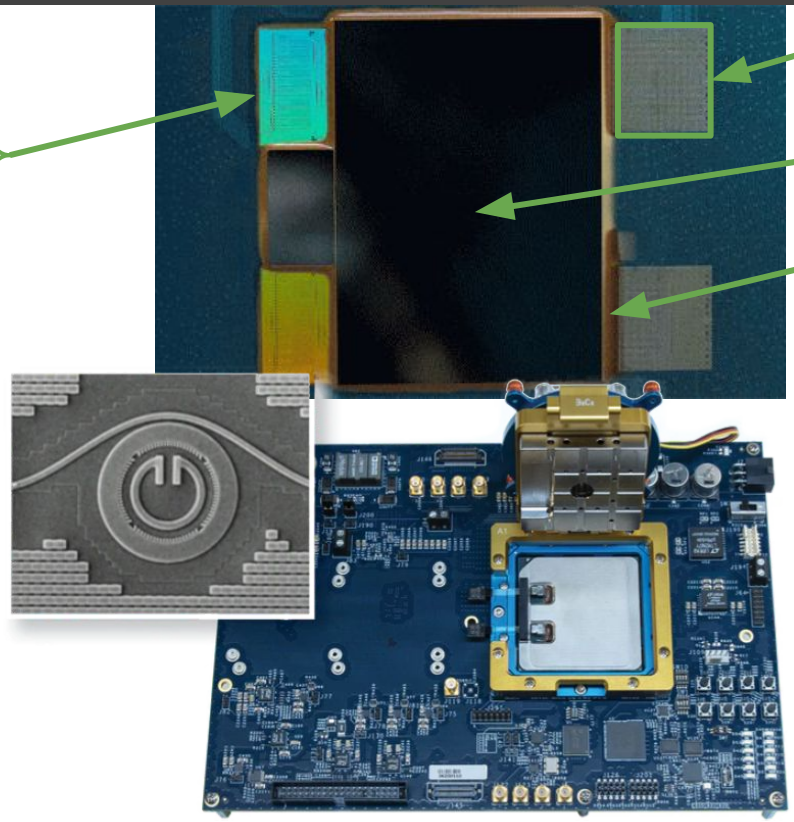


AIB Adopers

- Boeing
- Intrinsix
- Synopsys
- Intel
- Lockheed Martin
- Sandia
- Jariet
- NCSU
- U. of Michigan
- Ayar Labs

(2019) CHIPS HI Win #1, Photonic Interconnect

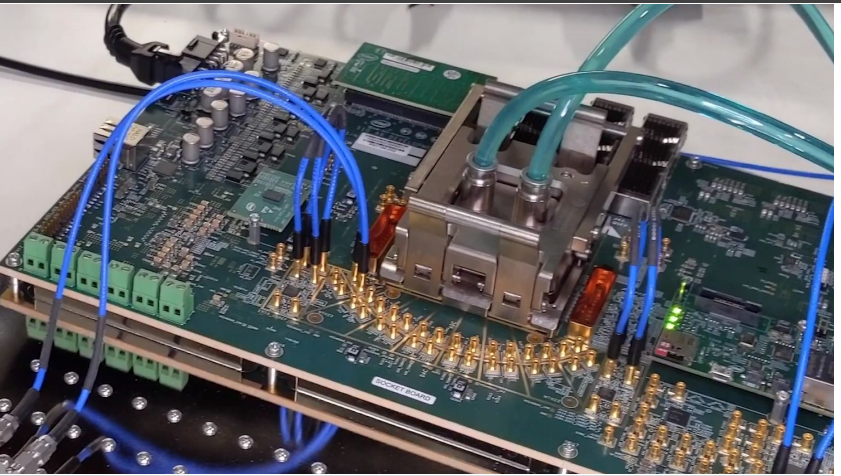


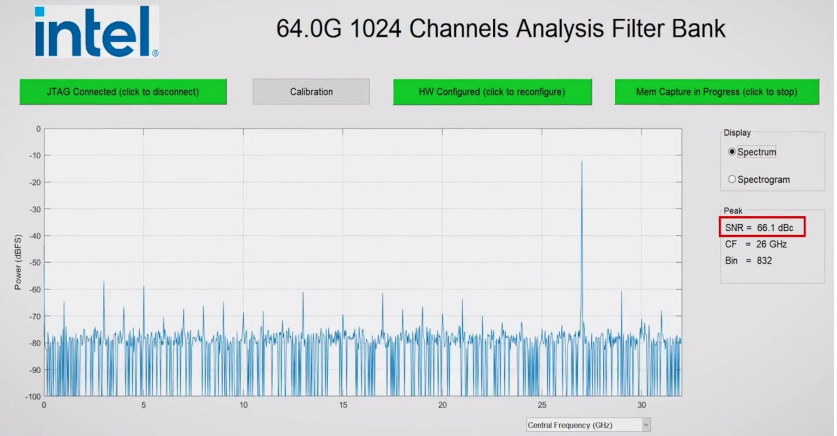


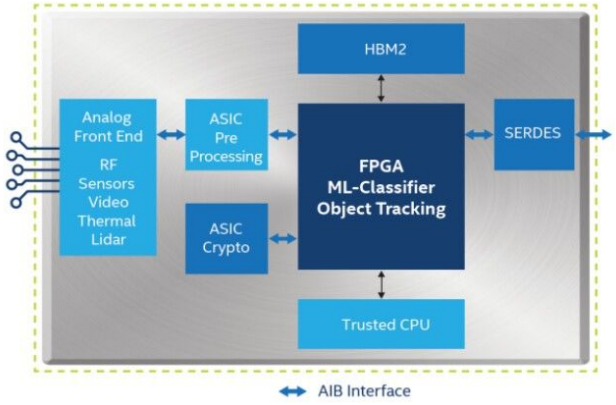
	Value
Bandwidth	1 Tbps
Height	8.86mm
Width	5.5mm
Latency	10ns
Reach	2km
Efficiency	<5pJ/bit

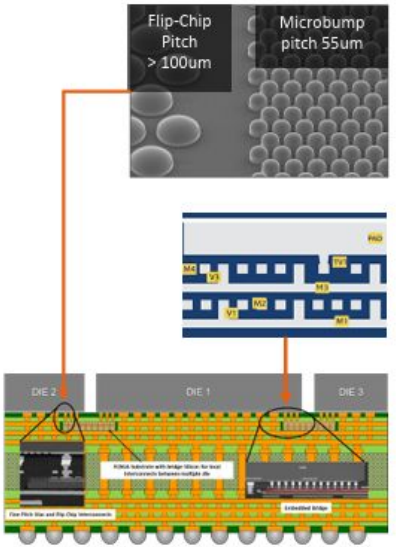
[REF] Wade (2019), “A Chiplet Technology for Low-Power, High-Bandwidth in-Package Optical I/O”, Hot Chips

(2019) CHIPS HI Win #2, Mixed Signal FPGAs





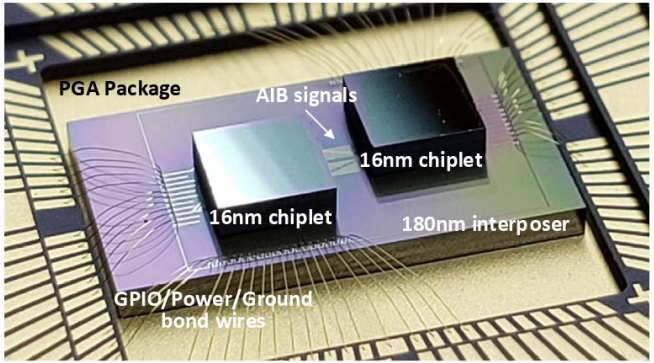




- 8 external silicon partners
- 4+ technology nodes
- 3 FPGA families
- 3 data converter chiplets
- 2 ASIC compete chiplets
- 9 serdes/optical IO chiplets

[REF] Shumarayev (2022), “Heterogenous Integration Enables FPGA Based Hardware Acceleration for RF Applications”, Hot Chips

(2019) CHIPS HI Win #3, Collaborative Innovation



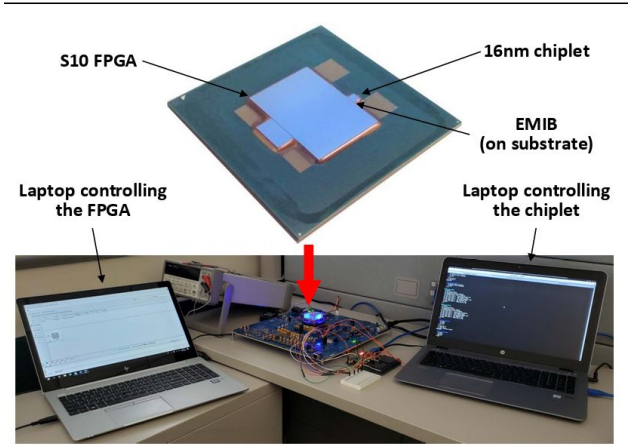
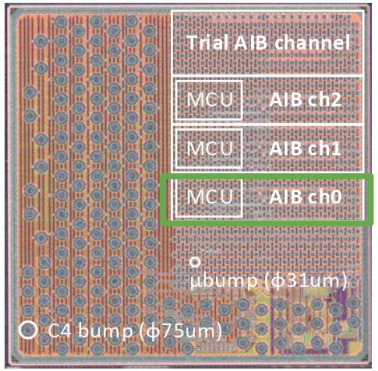
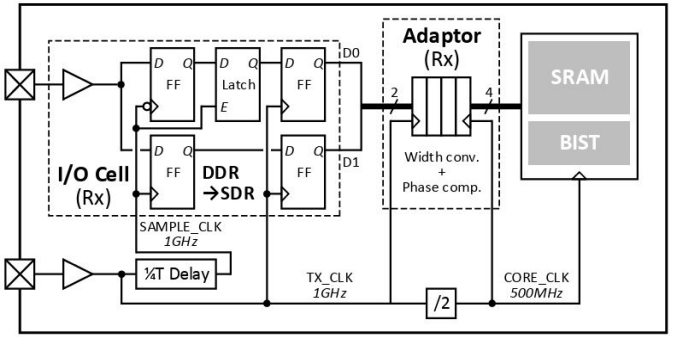


Fig. 5. A 16nm chiplet is integrated with an Intel Stratix 10 FPGA via EMIB on the package substrate.

	This Work
Technology	16nm FinFET
Voltage swing	0.9V
Bump pitch	55um
Chiplet carrier	Silicon interposer 3-layer / EMIB 4-layer
Reach	2mm
I/O size	203.2um^2/b
Data rate per pin	2Gb/s
Energy efficiency	0.83pJ/b
Shoreline BW density	256Gb/s/mm
Area BW density	614.4Gb/s/mm^2
Latency	4ns



Each AIB channel contains 96 signal and 42 power/ground µbumps, occupying 312.5µm × 1246.5µm



[REF] Liu, et al (2021), "A 256Gb/s/mm-shoreline AIB-Compatible 16nm FinFET CMOS Chiplet for 2.5D Integration with Stratix 10 FPGA on EMIB and Tiling on Silicon Interposer", IEE CICC

(2019) CHIP→SHIP→STEAMPIPE Transition

NEWS | Oct. 31, 2019

NSWC Crane leverages OTA to ensure that the U.S. Government has access to secure state-of-the-art design, assembly, packaging and test for state-of-the-art microelectronics

By NSWC Crane Corporate Communications

Andreas Olofsson, DARPA PM for the Common Heterogeneous Integration and IP Reuse Strategies (CHIPS) program said, "The future of computing hardware is specialized, heterogeneous and parallel."

CHIPS is a precursor for SHIP, and with the below stated goals it is serving as a transition partner to SHIP:

- Establish and demonstrate common interface standards
- Enable the assembly of systems from modular IP blocks built with these established standards
- Demonstrate reusability of the modular IP blocks via rapid design iteration

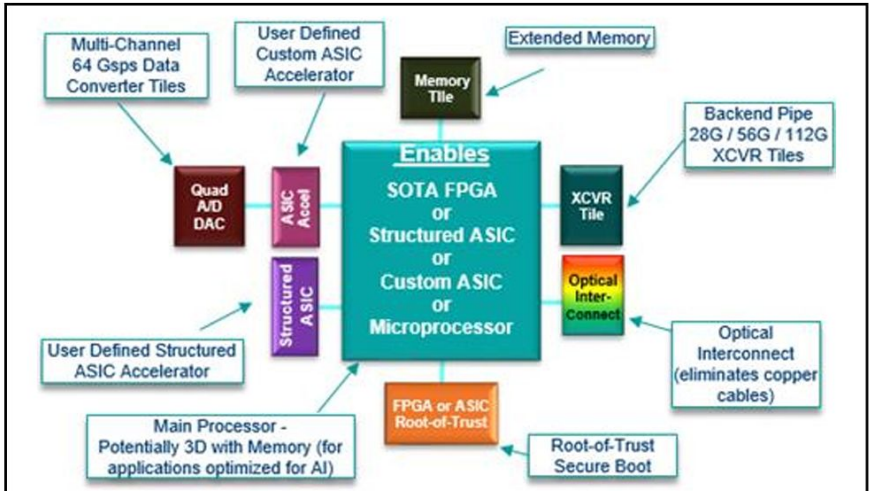
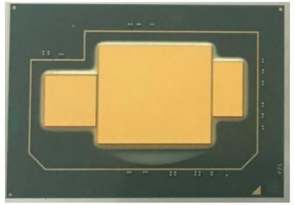
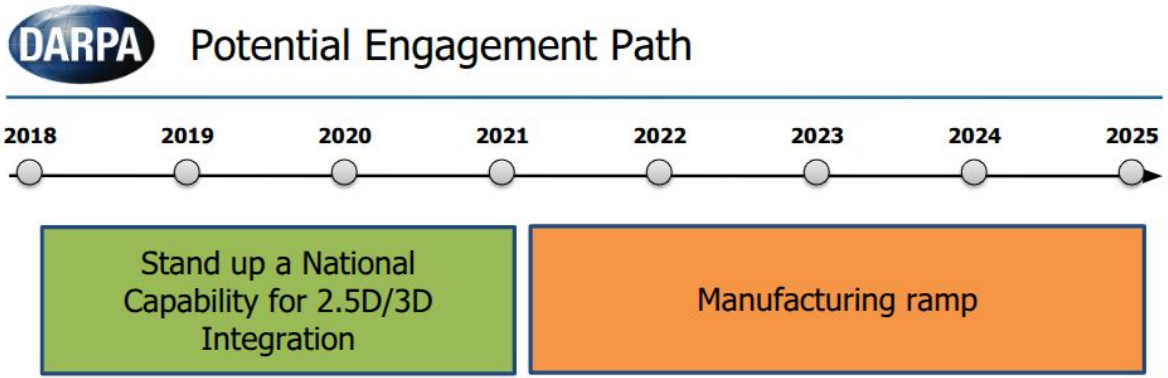


Figure 1. Notional Heterogeneous Integration Example



[REF] Shenoy et al (2023), "DoD Microelectronics: Heterogeneous Integration with Compound Semiconductors and Photonics", MANTECH

(2019) DARPA CHIPS 2.0 Workshop



- Commercial **on-shore** manufacturing
  - (See previous slide)
  - Si interposer w/ TSVs
  - Organic package substrates
  - Copper bumping (<=55 µm)
  - C4 bumping (150 µm)
  - 2.5D assembly
  - 3D assembly
  - Flip Chip Assembly
  - SOTA automation
- Assemble all silicon sources!
- Turnkey model

- "MOSIS for 2.5D"
- Agile PDK development
- Yield ramping
- Manufacturing cost optimization
- NPI cost optimization "zero" target
- Long Term Goals:
  - ~$20 turnkey packaging cost
  - 2 week assembly turn
  - Standard fab turns
  - Zero email order

It has been 6 years… how long until we have this in place?!

(2019) My DARPA 2025 Predictions

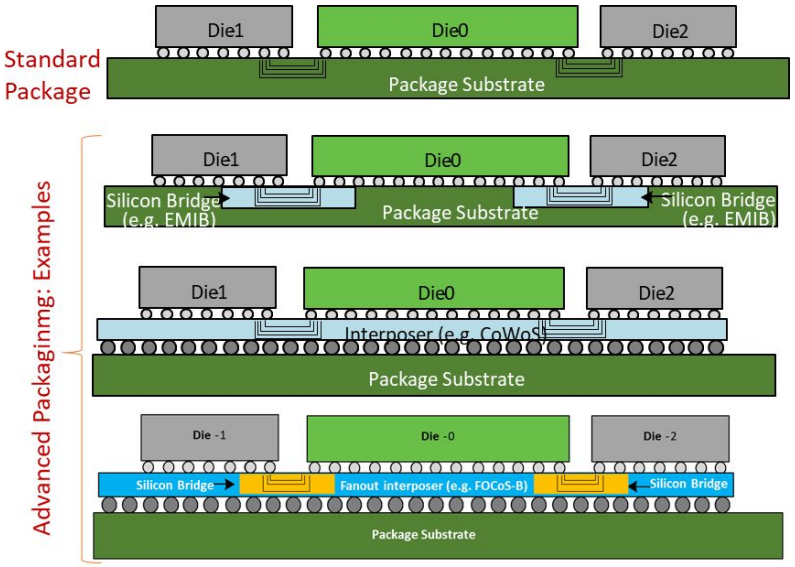


Conclusion: My 2025 Predictions…

- We will have no human in the loop general purpose silicon compiler (RTL/schematic →GDSII)
- We will experience FOSS "GCC/LLVM" for ASIC and FPGA design
- Domain specific compilers sitting on top of the silicon compiler will proliferate
- PCBs will be designed using programming languages, not schematic entry tools
- You will be able to download production quality analog & digital FOSS IP
- Building heterogeneous System-In-Package will be ~~as easy as~~ easier than PCB design
- ML ASICs will be ubiquitous
- All major system companies will design their own silicon
- Consumers will order custom "N=1" silicon

Distribution Statement "A" (Approved for Public Release, Distribution Unlimited)

(2022) UCIe Standard

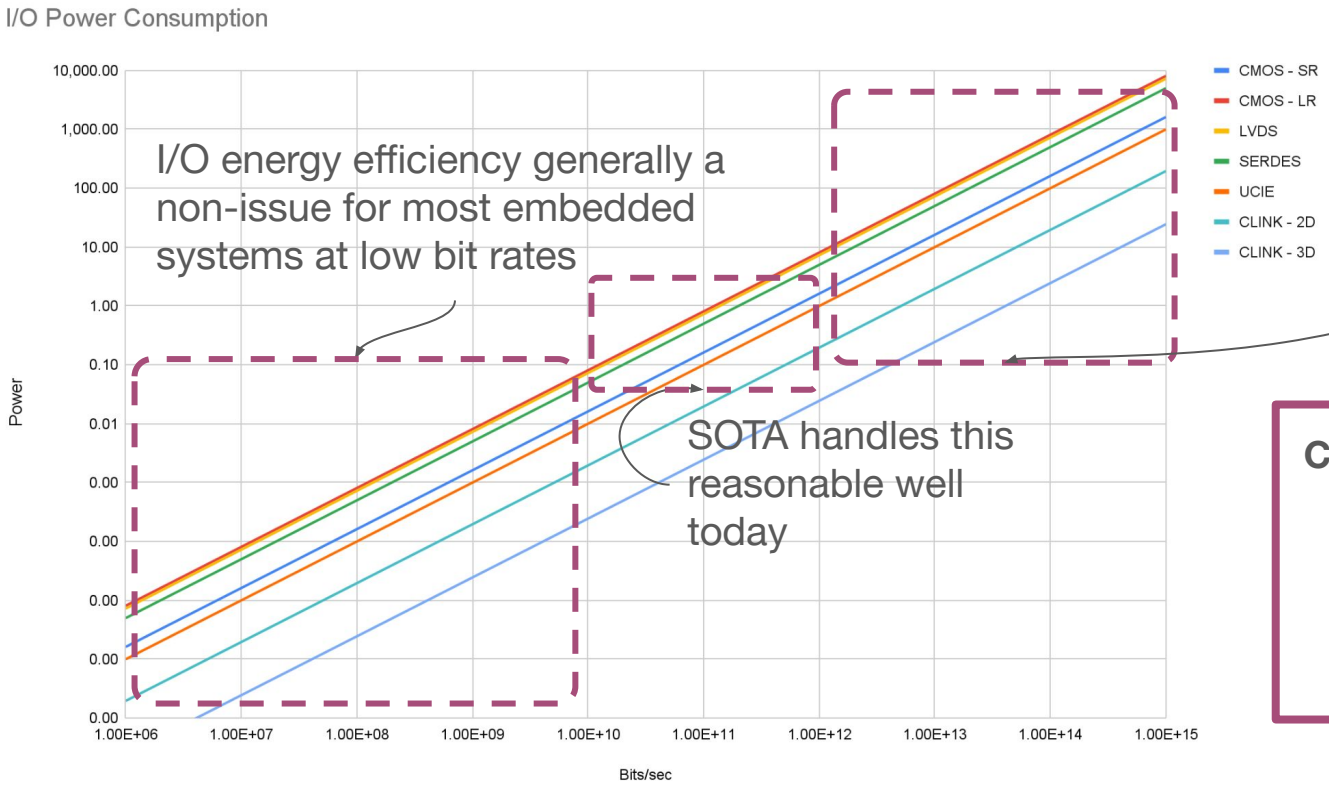


(b. Packaging Options: 2D and 2.5D)

Characteristics / KPIs	Standard Package	Advanced Package
Characteristics		
Data Rate (GT/s)	4, 8, 12, 16, 24, 32	
Width (each cluster)	16	64
Bump Pitch (um)	100 – 130	25 - 55
Channel Reach (mm)	<= 25	<=2
Target for Key Metrics		
B/W Shoreline (GB/s/mm)	28 – 224	165 – 1317
B/W Density (GB/s/mm^2)	22-125	188-1350
Power Efficiency target (pJ/b)	0.5	0.25
Low-power entry/exit	0.5ns <=16G, 0.5-1ns >=24G	
Latency (Tx + Rx)	< 2ns	
Reliability (FIT)	0 < FIT (Failure In Time) << 1	

TLDR: Big, fragmented, complex, expensive, not composable…but will likely find sockets in datacenter. What about everyone else?

(2025) Datacenter consuming all the chiplet oxygen



Driver of AI and datacenter interest in chiplets. Super high I/O bandwidths (>>1Tbps)

Chiplet Focus:

- 7nm and below
- > 1TBps BW
- Chiplet area > 64m^2



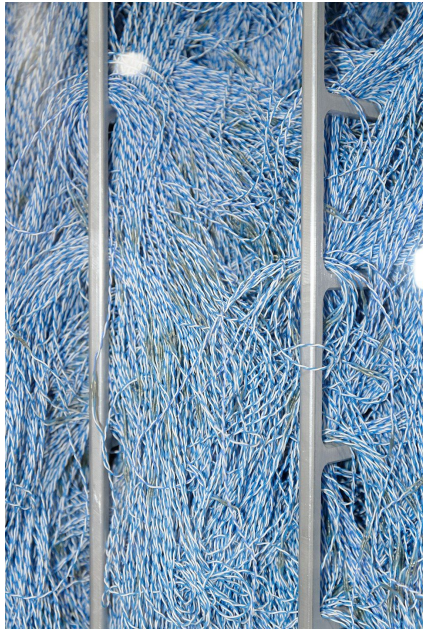
Zero ASIC Composable Chiplet Journey

(2020 - present)

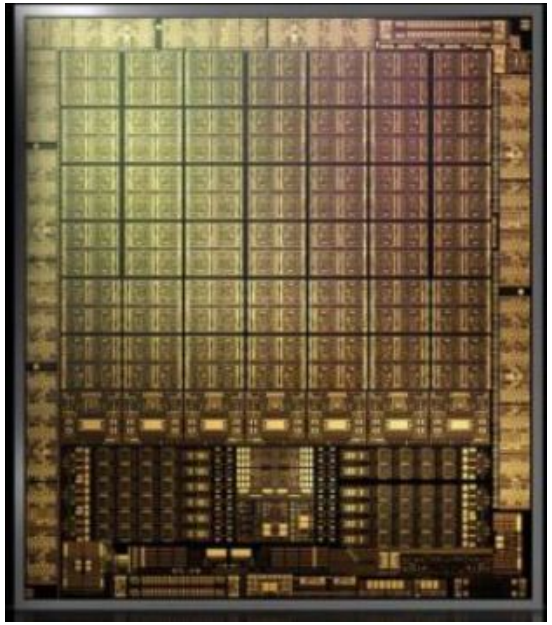
(2020) Heilmeier Questions

1. **What are we proposing?** *”LEGO for chiplets”*
2. **How is it done today?** Tower of Babel of bespoke chiplets
3. **What is new in our approach? A system of composable chiplets**
4. **Why does it matter?** Potentially cuts design time and cost by a factor of **100**
5. **What are the risks?** Disrupting 50 years of Moore’s law inertia
6. **How much will it cost?** $100M - $1B
7. **How long will it take?** 5 years
8. **What are key milestones?** First viable composable chiplet based system

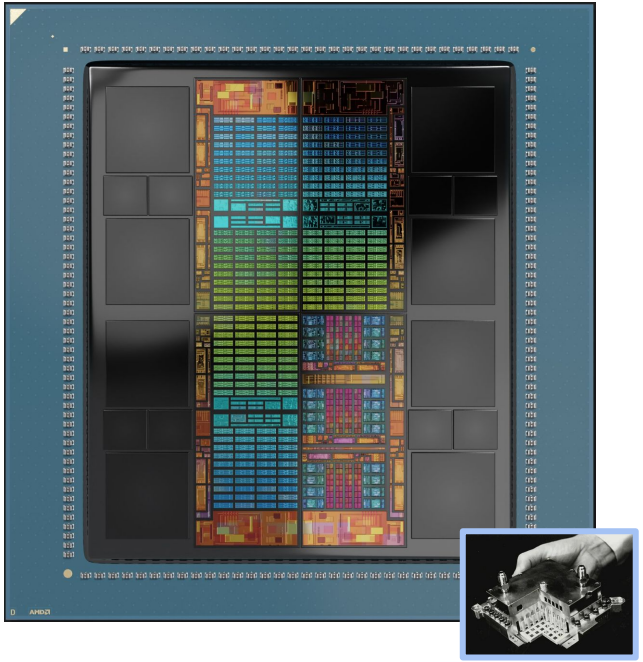
Three Er(r)a(r)s of Chip Design



Discrete Era
Tyranny of Wires
(1940 – present)

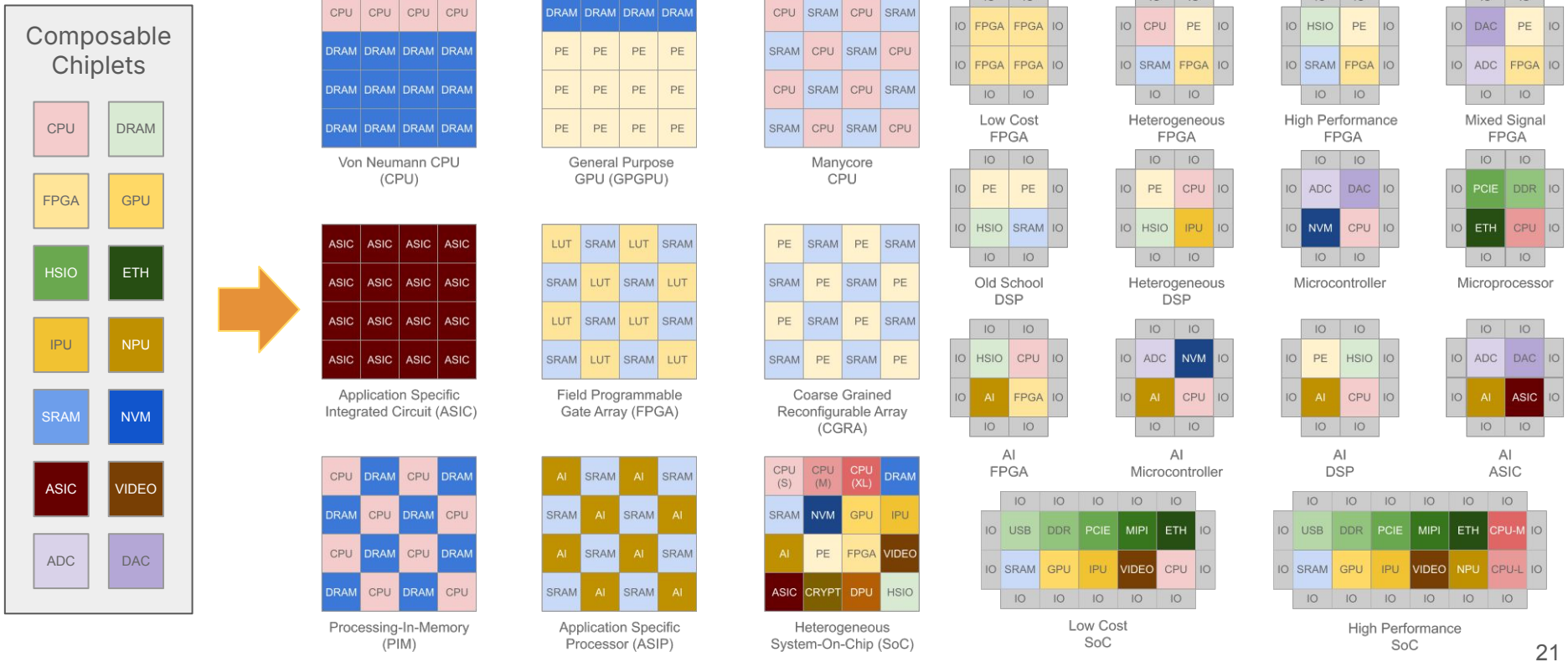


Monolithic Era
$1-10B Per Generation
(1960 – present)

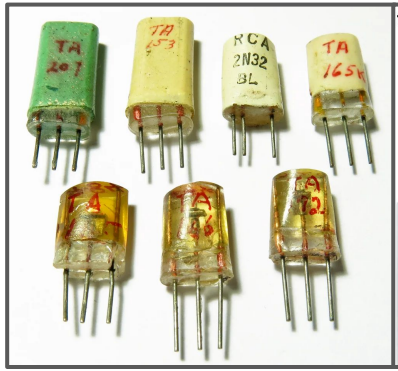


Chiplet Era
Private Bespoke Islands
(1980 - present)

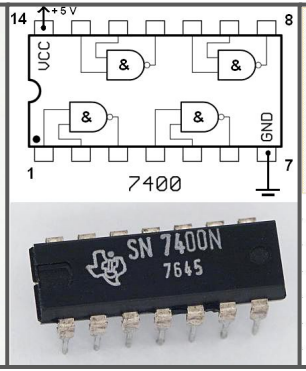
Could we build *“Amino acids for silicon systems”*



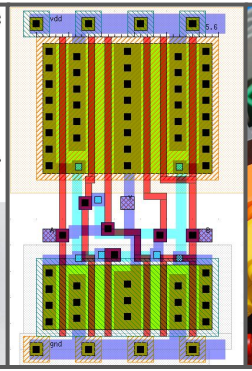
Composable Hardware Inspiration



Transistors



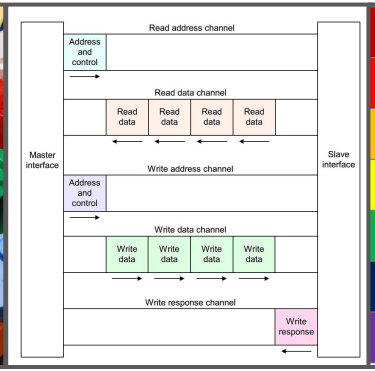
TTL Logic



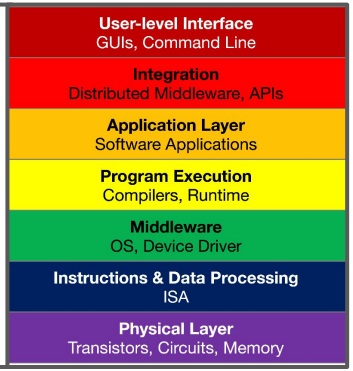
Logic Cells



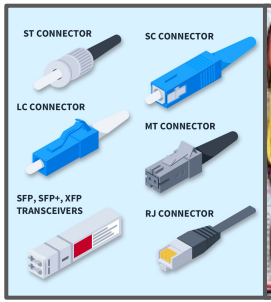
LEGO® Bricks



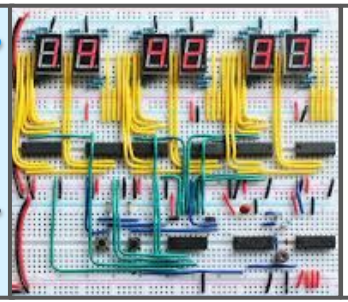
Amba IP



CPU Stack



Ethernet



Breadboard



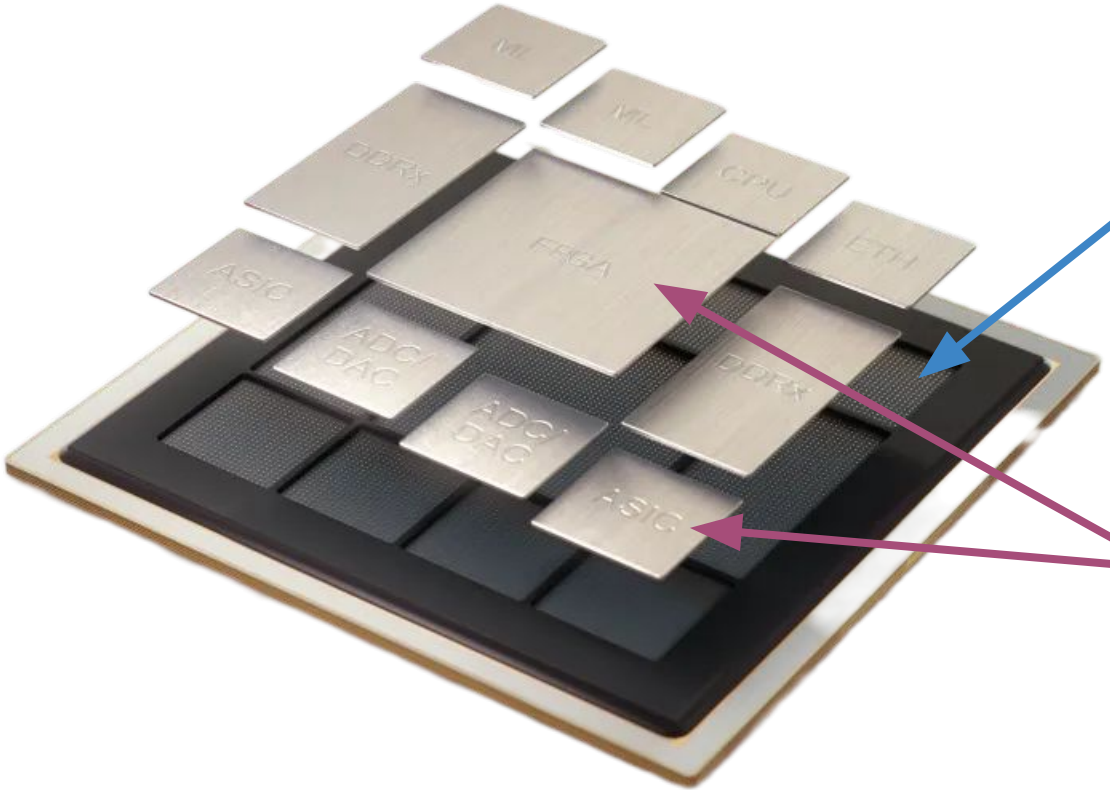
JEDEC DRAM

Composable hardware systems can be **effectively** constructed by connecting together independently developed modular and reusable components.

Key Composable Chiplet Optimization Questions

Question	Range	Conclusion
1. Mechanical Structure	2D, 2.5D, 3D	3D
2. Substrate Technology	Organic, glass, Si (active/passive),...	Active Silicon
3. Chiplet Types	FPGA, CPU, ML, SRAM, DRAM,...	Many…
4. Chiplet sizes	1 mm^2 to 858 mm^2	Discrete grid
5. Interconnect Pitch	1 um to 150 um	45 um →8um →4um →…
6. Standard	UCIe, BOW, AIB, HBM, …	CLINK + EBRICK + UMI

Zero ASIC's Composable Chiplet Approach



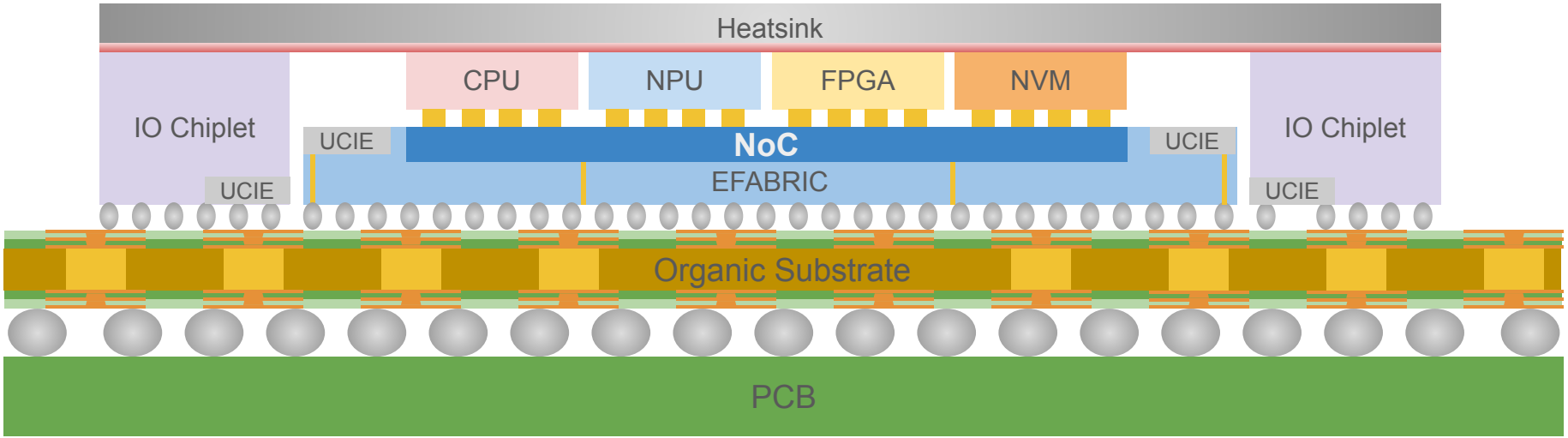
EFABRIC

- Active silicon substrate
- Fixed mechanical grid connections
- Built in NoC, clocking, management
- Shared memory architecture
- 3D chiplet links
- Scale out 2D I/O

EBRICKS

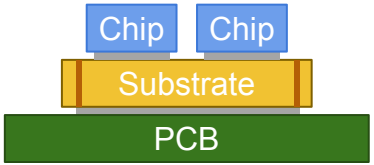
- Discretized chiplet sizes
- 3D chiplet point-to-point links
- CPU, FPGA, NPU, etc, …
- 100% interchangeable/swappable
- Rotationally symmetric footprints
- Rigid specification (aka like ethernet)

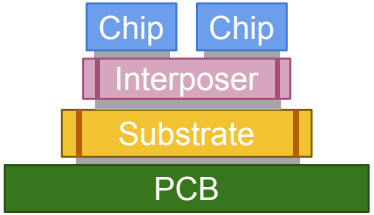
EFABRIC Cross Section (v2)

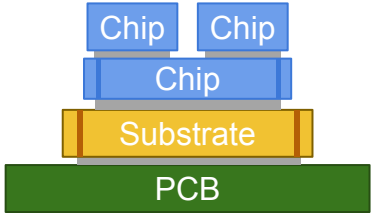


- Optimized for manufacturability, performance, and supply chain security
- 45um 3D bumps, 110um I/O bumps, 100um chiplet spacing

Q{1,2}: Mechanical Topology

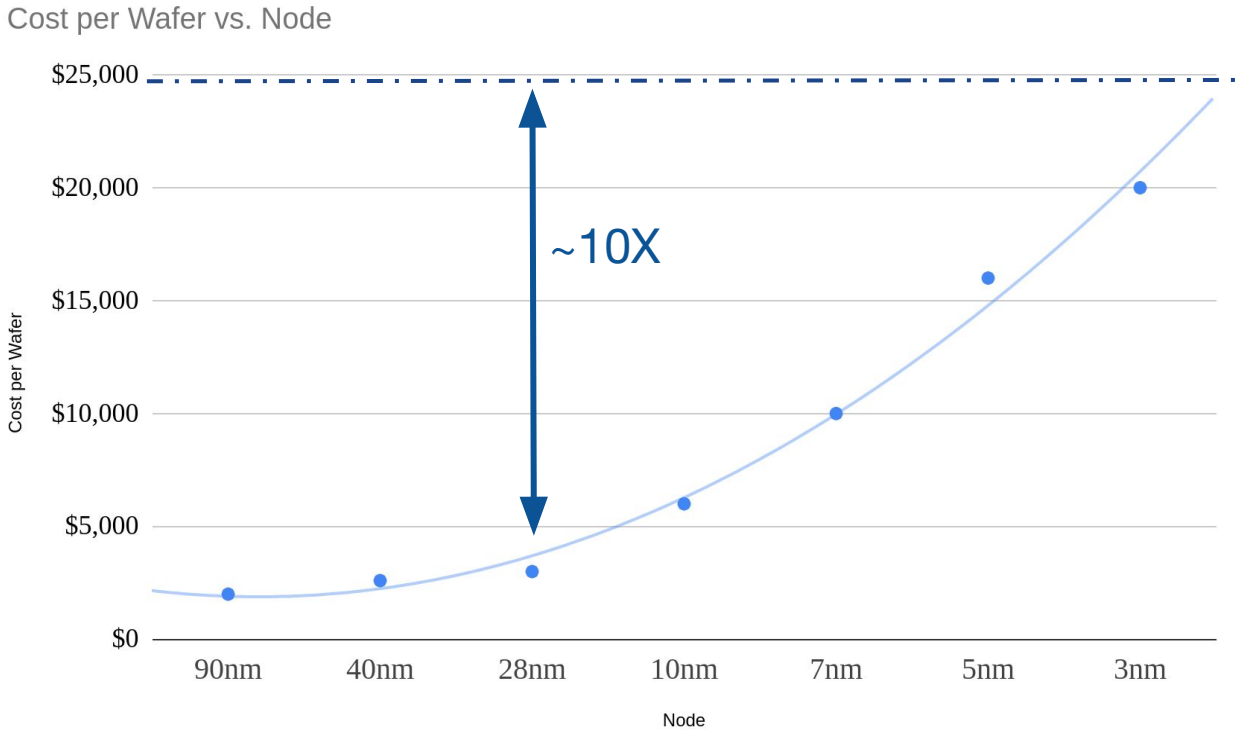




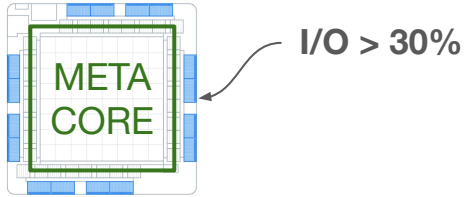


	2D	2.5D	3D
Wire Length	1000um - 5000 um	1000um - 5000 um	< 100 um
Wire Density	50 wires/mm/layer	500 wires/mm/layer	500 - 10,000 wires / mm^2
Cost	Low	Medium	**Medium**
Mfg Risk	Low	High	**Medium**

A{1,2}: Debunking Myth of Expensive Si



- New silicon is expensive
- Old silicon is inexpensive
- Make interposers in old silicon. 🤯
- Wafer based backend is inexpensive
- Minimize total cost by splitting out non-scalable I/O to old process.
- Benefits of active SI outweigh cost
- **Caveat**: chiplet interface must be small



Q&A 3: Selecting Chiplet Types

	Snapdragon 8	NXP MX8+	NVDA Orin	AMD Zynq
CPU	✅	✅	✅	✅
DDRx	✅	✅	✅	✅
NPU	✅	✅	✅	
GPU	✅	✅	✅	✅
DSP	✅	✅	✅	✅
FPGA				✅
Serdes			✅	✅
Other	✅	✅	✅	✅

Studies

#1: BoM based study
#2: SoC tear down (simplified)
#3: Clean sheet design
#4: Digikey web scraping
#5: Customer Survey

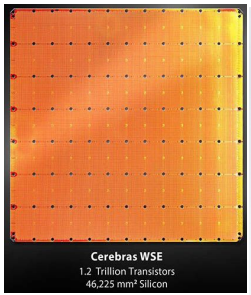
MSB Choices Obvious

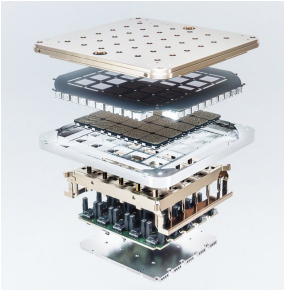
CPU, DDR, FPGA, SERDES, NPU

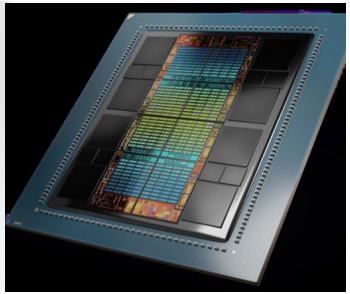
LSB Choices Trickier

…

Q4: What Is the optimal chiplet size?







I/O communication costs are prohibitive, so maximum die are optimal for large problems

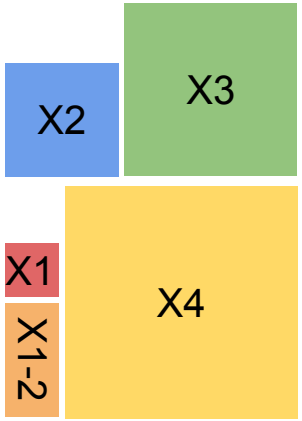
Conclusion:
Composability favors small dies, HPC favors large dies. No optimal chiplet size so we need to support multiple sizes.

Lib size (n)	5	10	5	10
Fabric Area (A)	100 mm^2	100mm^2	858mm^2	858mm^2
Chiplet (C)	**Composability (n^(A/C))**			
1	7.89E+69	1.00E+100	#NUM!	#NUM!
4	2.98E+17	1.00E+25	3.80E+149	1.00E+214
9	4.88E+07	1.00E+11	2.52E+66	1.00E+95
16	1.56E+04	1.00E+06	1.11E+37	1.00E+53
25	625	10,000	5.82E+23	1.00E+34
36	25	100	1.19E+16	1.00E+23
49	25	100	7.63E+11	1.00E+17
64	5	10	1.22E+09	1.00E+13
81	5	10	9.77E+06	1.00E+10
100	5	10	3.91E+05	1.00E+08

Composability (“solution diversity”) achieved via small dies and large substrates.

A4: Standardized Discretized Chiplet Grids

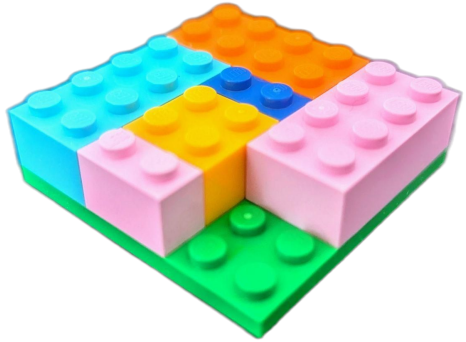
Name	Size
X1	0.95mm
X2	2mm
X3	3.05mm
X4	4.1mm
X5	5.15mm

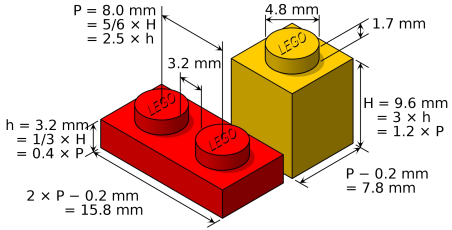


Key Considerations:

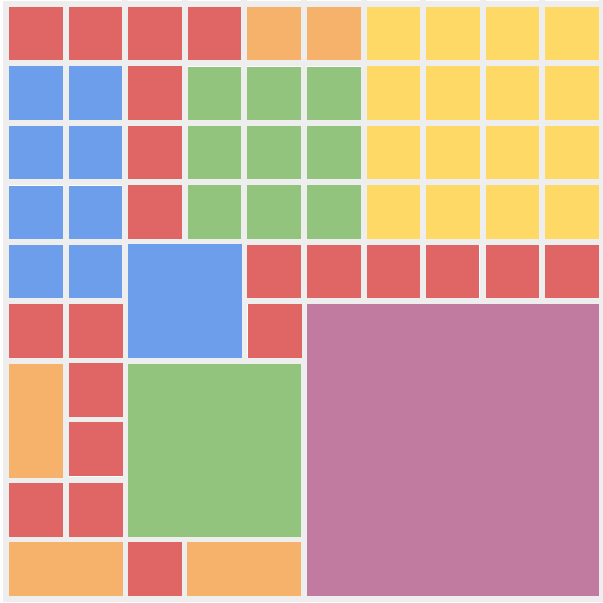
- Cost/density of 18A/N3 silicon
- 100Mtr/mm^2
- 100um safe chiplet spacing
- Minimum handling size
- Composability
- “Forever standard”

Inspiration



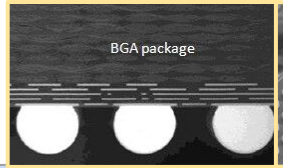
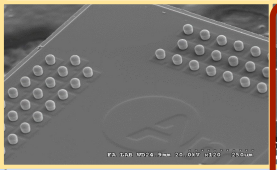
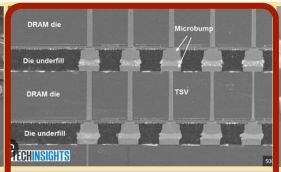
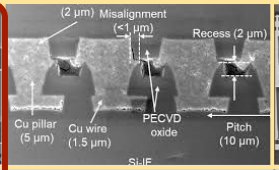
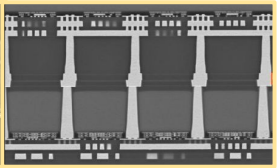


Lego Brick Standard Unchanged Since 1958!



Fixed Forever Chip Grid!

Q&A5: 3D Chiplet Interconnect Pitch

					
Pitch	**150um**	**110um**	**45um**	**10um**	**5um**
Pins/mm^2	44	82	493	10,000	40,000
Interface	Solder Ball	Cu+SnAg	Cu+SnAg	Cu	Cu
Assembly	Reflow	Reflow	Reflow	TCB	Hybrid
Cost	Low	Low	Medium	High	High
Tech Risk	Low	Low	Medium	High	High
REF	OSAT	OSAT	HBM	UCLA	AMD

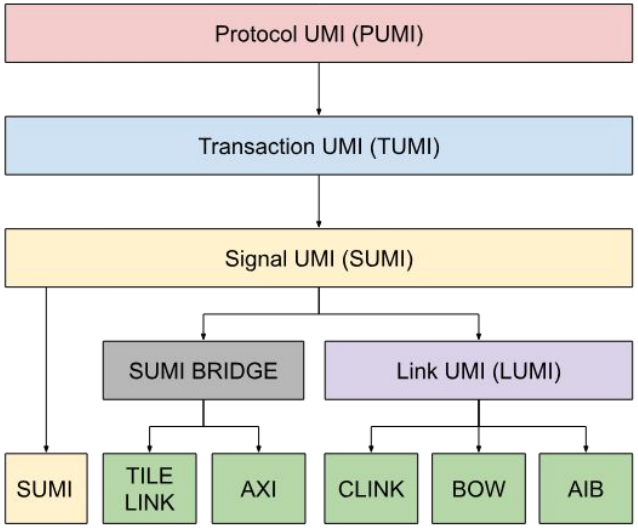
No right answer, but many wrong answers…

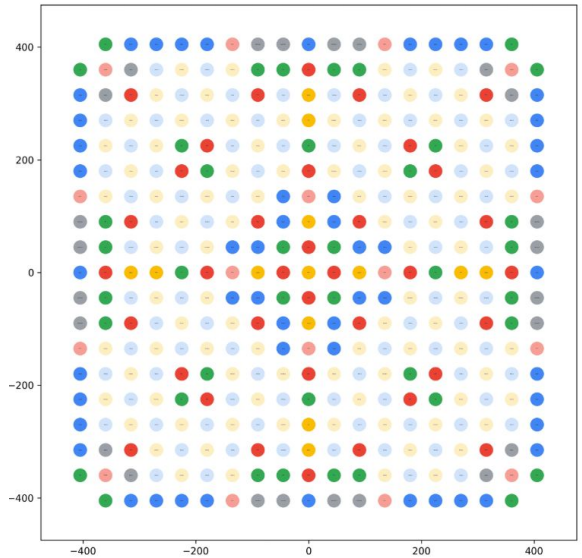
Q6: 3D Chiplet Standard

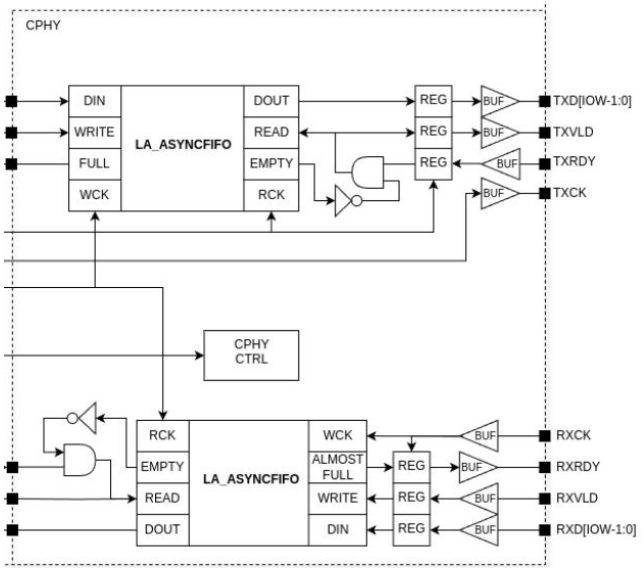
	AIB	BOW	UCIE
Adoption Rate	Abandoned?	Unclear	High
Electrical standard	Yes	Yes	Yes
Footprint standard	No	No	No
Protocol Standard	No	No	No
3D Standard	No	No	Yes
Symmetrical	No	No	No
Suitable as AXI replacement	No	No	No

Existing chiplet standards don’t support composability.

A6: A full stack 3D chiplet standard

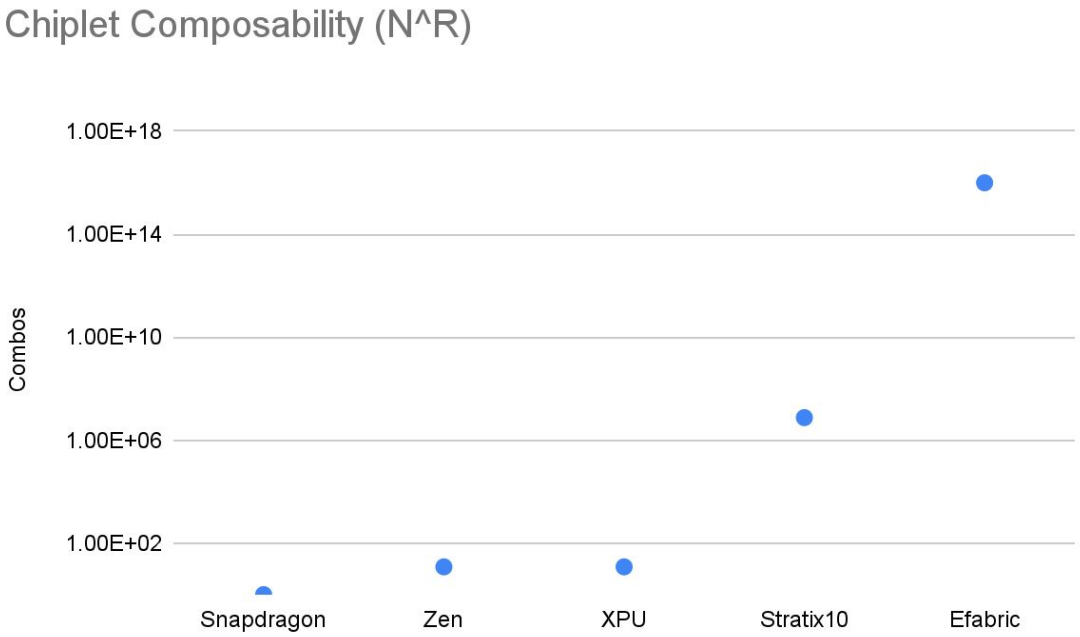


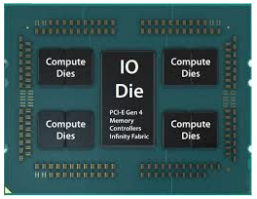




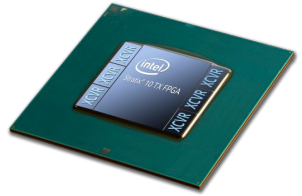
UMI Protocol	EBRICK Footprint	CLINK Electrical
Memory mapped packets	64b datapath	Source synchronous
Latency Insensitive	Rotational Symmetry	8b - 1024b
github.com/zeroasiccorp/umi	Analog, multi-power, passthrough	0.04mm2 in ASAP7 (512b)

EFABRIC: Composability Comparison

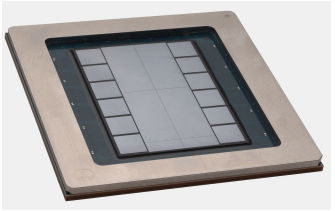




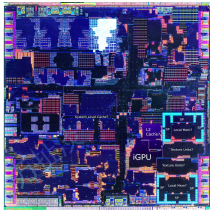
Zen



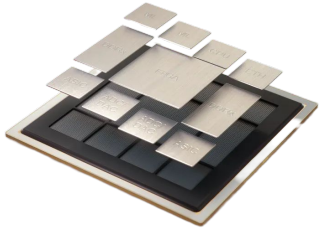
Stratix 10



XPU



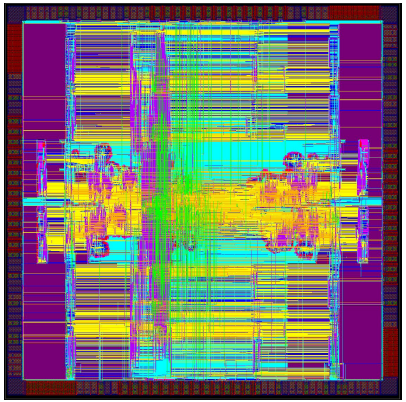
Snapdragon

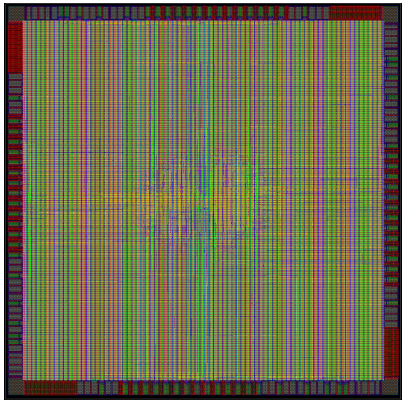


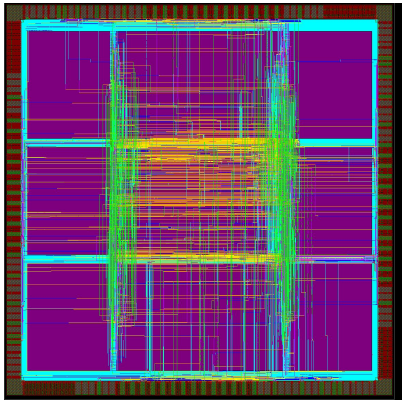
Efabric

[1] https://www.amd.com/en/products/processors/server/epyc/4th-generation-9004-and-8004-series.html
[2] https://www.intel.com/content/www/us/en/products/details/fpga/stratix/10.html
[3] https://www.broadcom.com/info/ai/3point5d
[4] https://chipsandcheese.com/p/inside-the-snapdragon-855s-igpu
[5] Zero ASIC, N = 10 (library size), R = 16 (number of sockets)

EBRICK: Composable Chiplet Prototypes



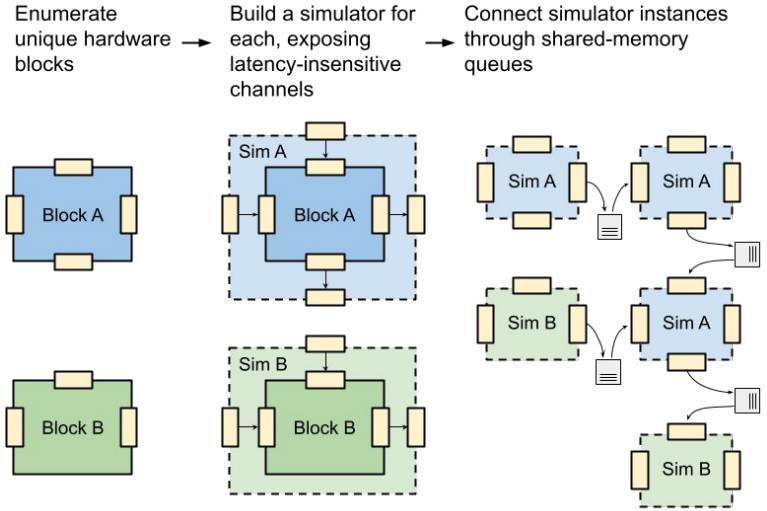




	GOTLAND	MAUI	KODIAK
PROCESS	12nm	12nm	12nm
STANDARD	EBRICK_2x2	EBRICK_2x2	EBRICK_2x2
TYPE	CPU	FPGA	MEMORY
SIZE	2 x 2 mm	2 x 2 mm	2 x 2 mm
METRIC	Quad Core RV64GC CPU	(now 2K LUTs/mm^2)*	3MB
DESIGNERS	2	2	2
WALL TIME	< 4 weeks	< 8 weeks	< 8 weeks
RUN TIME	< 24hrs	< 24hrs	< 24hrs

Switchboard: Chiplet Design Abstraction

- Heterogeneous simulation framework
- Latency insensitive protocol (ready/valid)
- Fast shared memory queues
- Supports RTL, FPGAs (HIL), SW models)
- UMI implementation
- Python bindings
- 10x faster than commercial emulators
- 1000X build time improvement over Verilator
- **Deployed in AWS**
  - 0.2us host latency
  - 4us host-fpga latency
- **Source**: github.com/zeroasiccorp/switchboard
- **Demo**: zeroasic.com/emulation



TIMING BREAKDOWN FOR THE MILLION-CORE SIMULATION

Name	Time	Percentage
Launch 250 ECS tasks	2m 30s	23%
Wait for ECS tasks to boot	1m 20s	12%
Run simulation	7m 4s	65%
Total	10m 54s	100%

S.Herbst, et al, Switchboard: An Open-Source Framework for Modular Simulation of Large Hardware Systems, arXiv preprint arXiv:2407.20537, Jul 2024

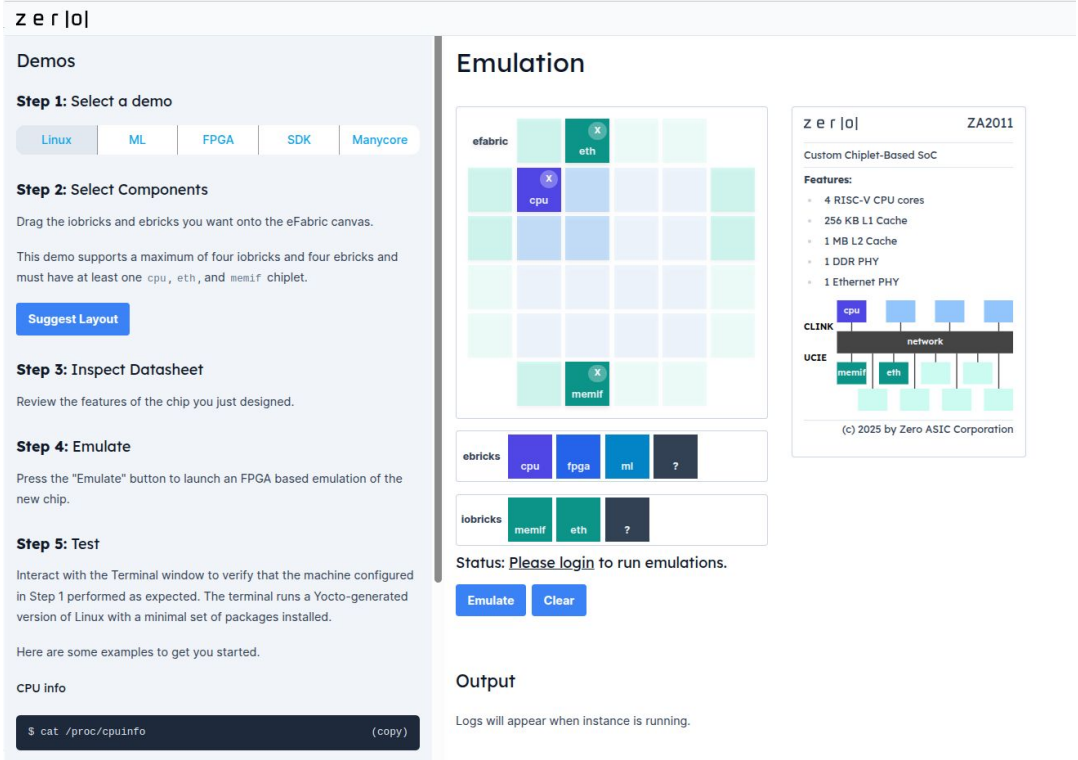
Emulator: Chiplet Digital Twin Demo

Removing Barriers:

- **No** EDA licensing
- **No** IP licensing
- **No** code
- **No** layout
- **No** mask costs
- **No** fabrication
- **No** installation

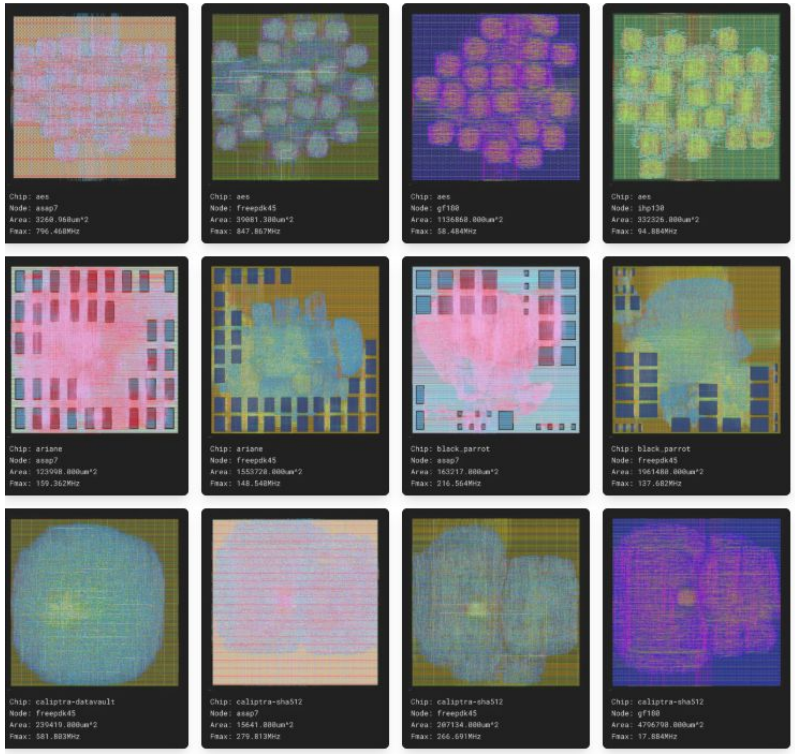
New Capabilities:

- White box validation
- Real time emulation



https://emulation.zeroasic.com/emulation

SiliconCompiler: Automated chiplet compilation



```
$ pip install siliconcompiler
$ sc heartbeat.v -remote
```

```
import siliconcompiler
chip = siliconcompiler.Chip('heartbeat')
chip.load_target('skywater130_demo')
chip.input('heartbeat.v')
chip.clock('clk', period=10)
chip.set('option','remote', True)
chip.run()
chip.summary()
chip.show()
```

PDKs: GF12LP, GF22FDX, SKY130, GF180
Tools: Yosys, Openroad, VPR, Verilator, Icarus, Xyce, GHDL, Slang, Klayout, Cadence, Synopsys, Siemens, and many more

https://github.com/siliconcompiler

A. Olofsson, et al. "Invited: A Distributed Approach to Silicon Compilation", DAC 2022,

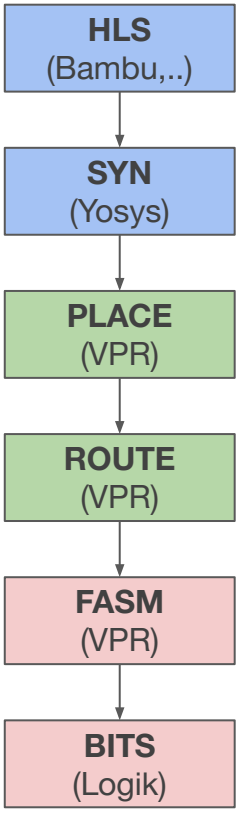
Platypus: Because we need a “RISC-V for FPGAs”

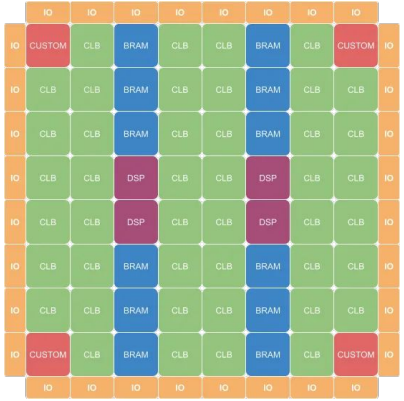
Zero ASIC · Mar 18, 2025

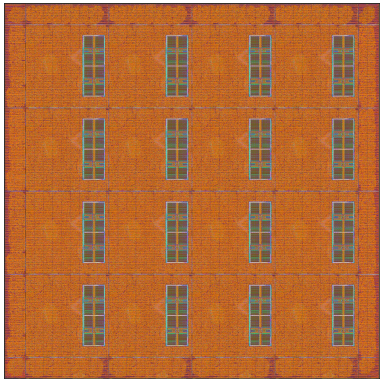
Zero ASIC launches world's first open standard eFPGA product

Cambridge, MA – March 18, 2025 – Zero ASIC, a U.S. semiconductor startup on a mission to democratize silicon, today announced Platypus™, the world's first open standard eFPGA product. Platypus addresses a long standing critical issue of FPGA obsolescence and vendor lock that has put critical infrastructure at risk.

- 100% open standardized FPGA architecture
- 100% open source FPGA bitstream format
- 100% open source FPGA development tools
- 2K LUTs in 1mm^2
- Support for BRAM and DSPs
- GF12LP process node (other nodes in development)
- OpenRoad based PNR implementation
- **Will become a standardized chiplet!!**
- https://github.com/siliconcompiler/logik
- https://github.com/siliconcompiler/logiklib/releases



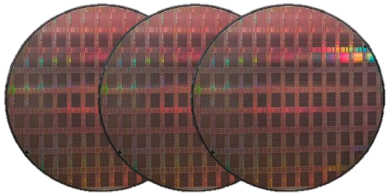




zerlol

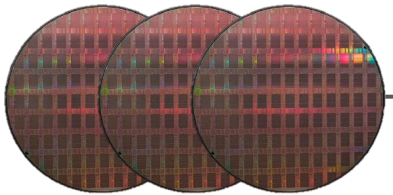
Predicting The Future of Chiplets

Lights Out Chiplet Assembly



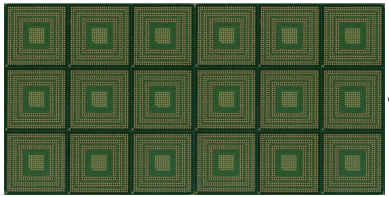
Chiplet Inventory

API



Interposer Inventory

API

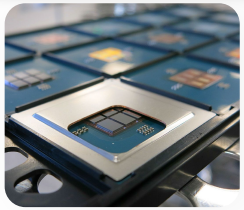


Substrate Inventory

API

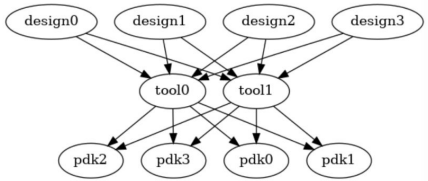
AUTOMATED MANUFACTURING PIPELINE



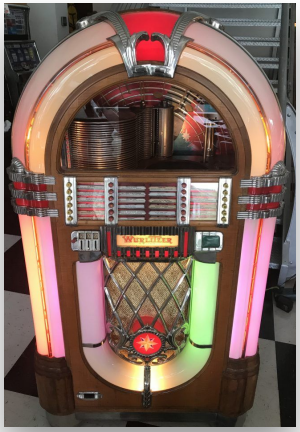


Standardized automation is the only way to fix the broken economics of low-volume high-mix manufacturing

New Era of Mechanical Configurability

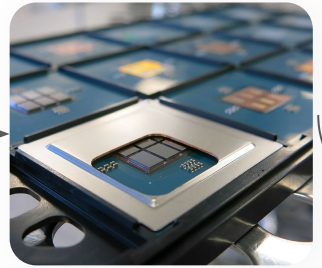


100% Automated
Silicon Compilers





100% Automated
System-In-Package



Modular
Device Library

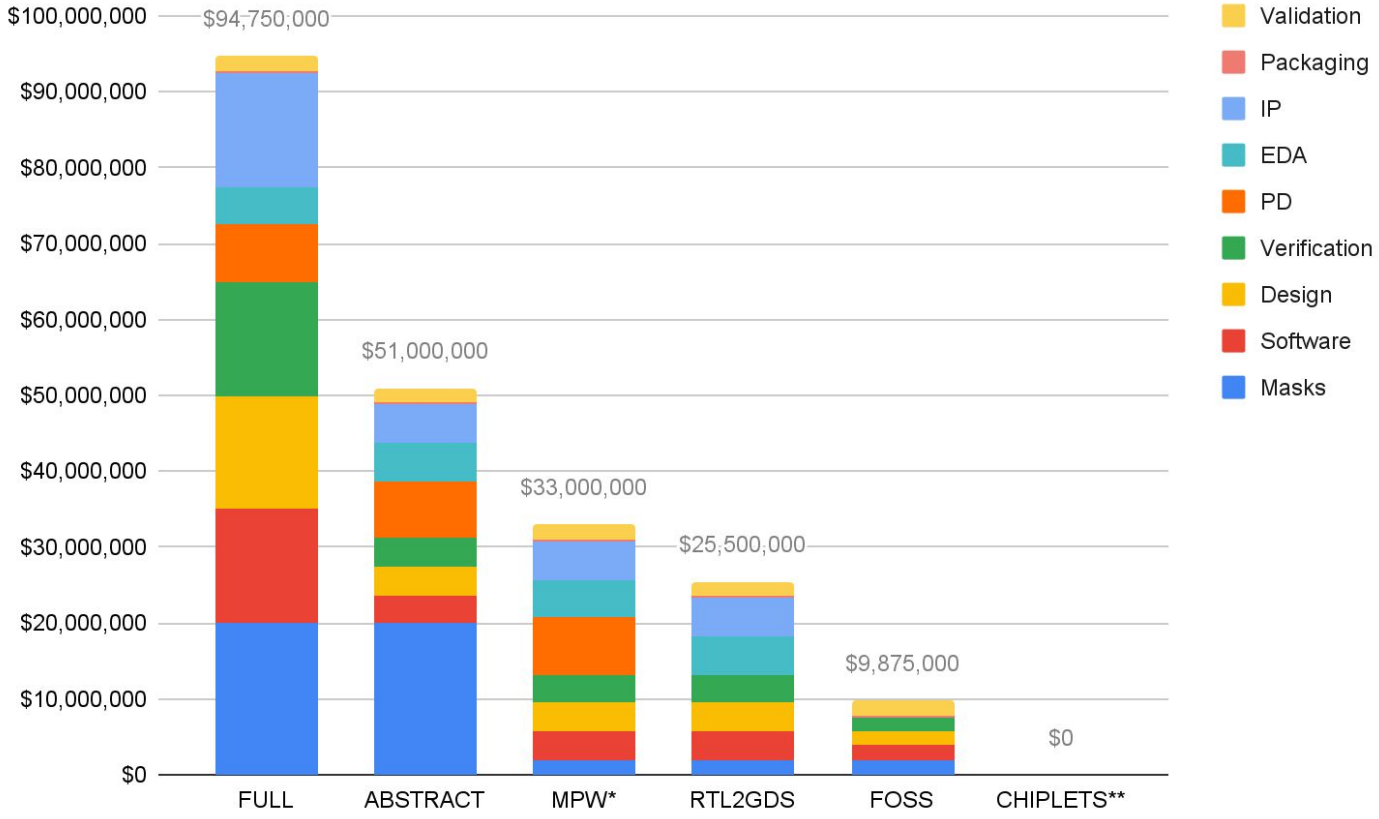
100% Automated Robotic
Reconfiguration

New Silicon	90 days
New Device	1 day
New Supercomputer	1 day

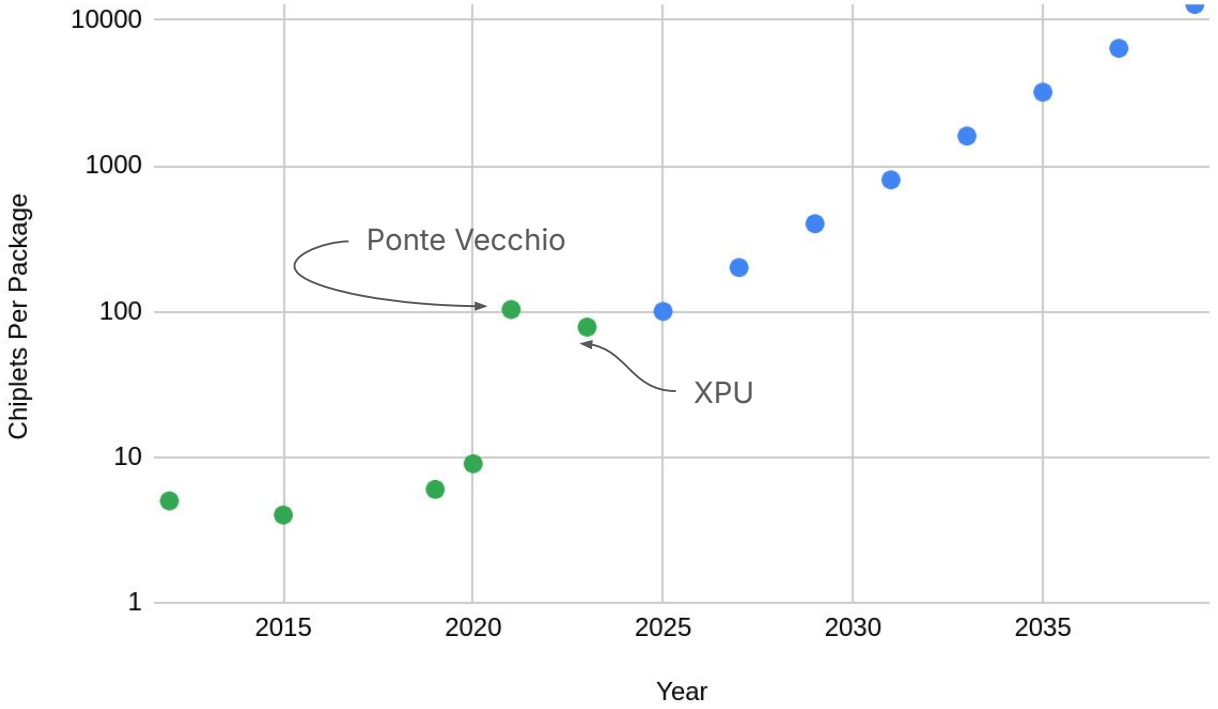




One Day We Will Spin SiPs at a cost of $1K in 24Hrs



Olofsson's Chiplet Roadmap



"The number of chiplets in a package will double every two years."

* See Stiglers' Law

Conclusion

“The most reliable way to predict the future is to create it.”

– Abe Lincoln

