ALL PROGRAMMABLE
ANY MEDIA
ANY MACHINE
5G
4K/8K
ANY STANDARD
ANY NETWORK
5G Wireless • SDN/NFV • Video/Vision • ADAS • Industrial IoT • Cloud Computing

XILINX
ALL PROGRAMMABLE™

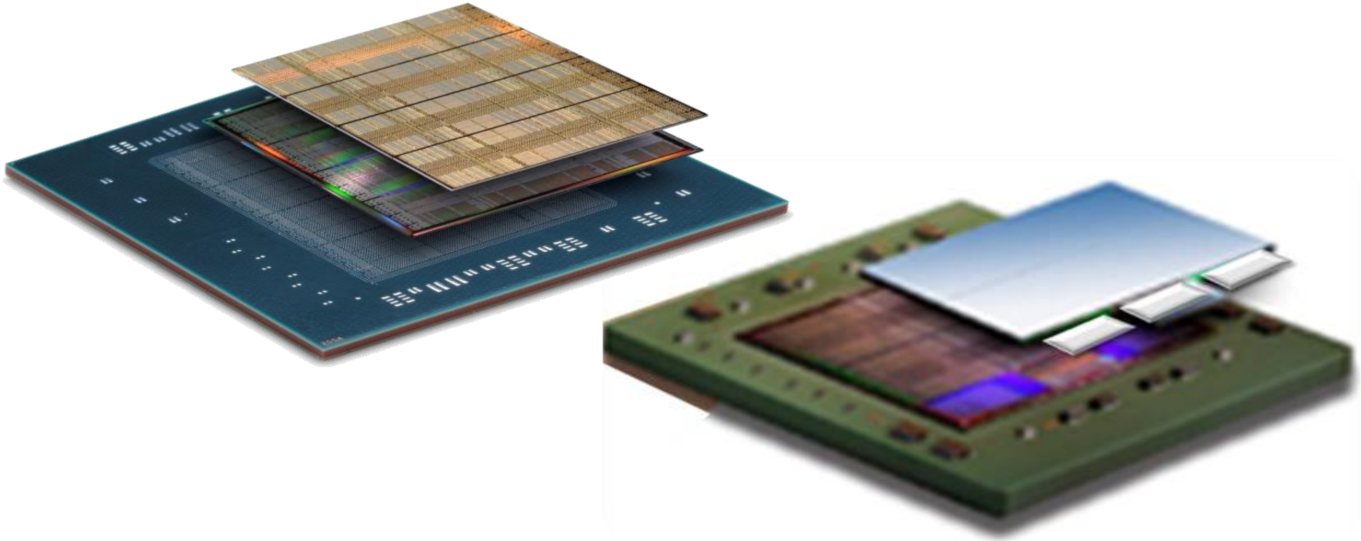
HBM Package Integration: Technology Trends, Challenges and Applications

Suresh Ramalingam

Hot Chips
Aug 21, 2016

Agenda

- Motivation
- HBM Packaging Options
- Interposer Design
- Supply Chain
- Application and Challenges
- Summary



Stacked ***S***ilicon ***I***nterconnect ***T***echnology Refers to Xilinx 3D solutions

FPGA’s in the Data Center Today

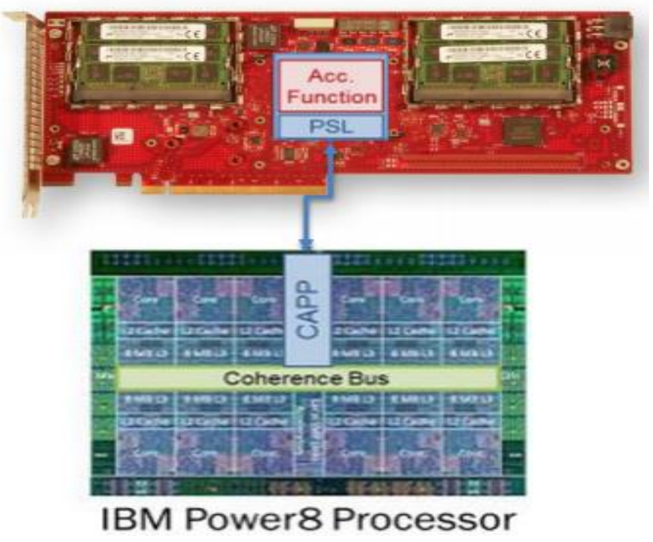
- **The Accelerator (FPGA or GPU) is used to offload only certain tasks**
  - These tasks are called “Workloads”, and FPGA’s are well suited for many workloads
  - Note: Accelerators <u>don’t</u> replace the CPU!

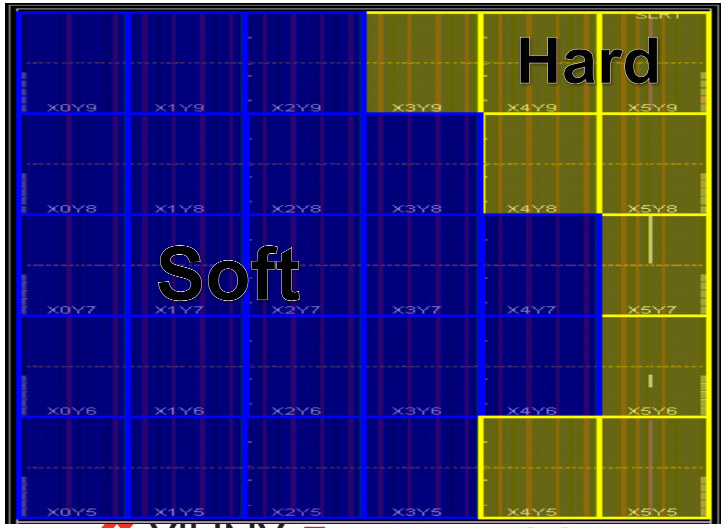
- **Hard and Soft is the basic approach**
  - **Hard** is the IO, Memory and PCIe interfaces
    - Does not change
  - **Soft** is the workload being accelerated
    - Is configured on the fly using P.R.

- **API’s are run on the CPU to reprogram the FPGA to accelerate the workload as needed.**
  - Average P.R. happens every 15 minutes!

- **Acceleration requires lots of memory BW**

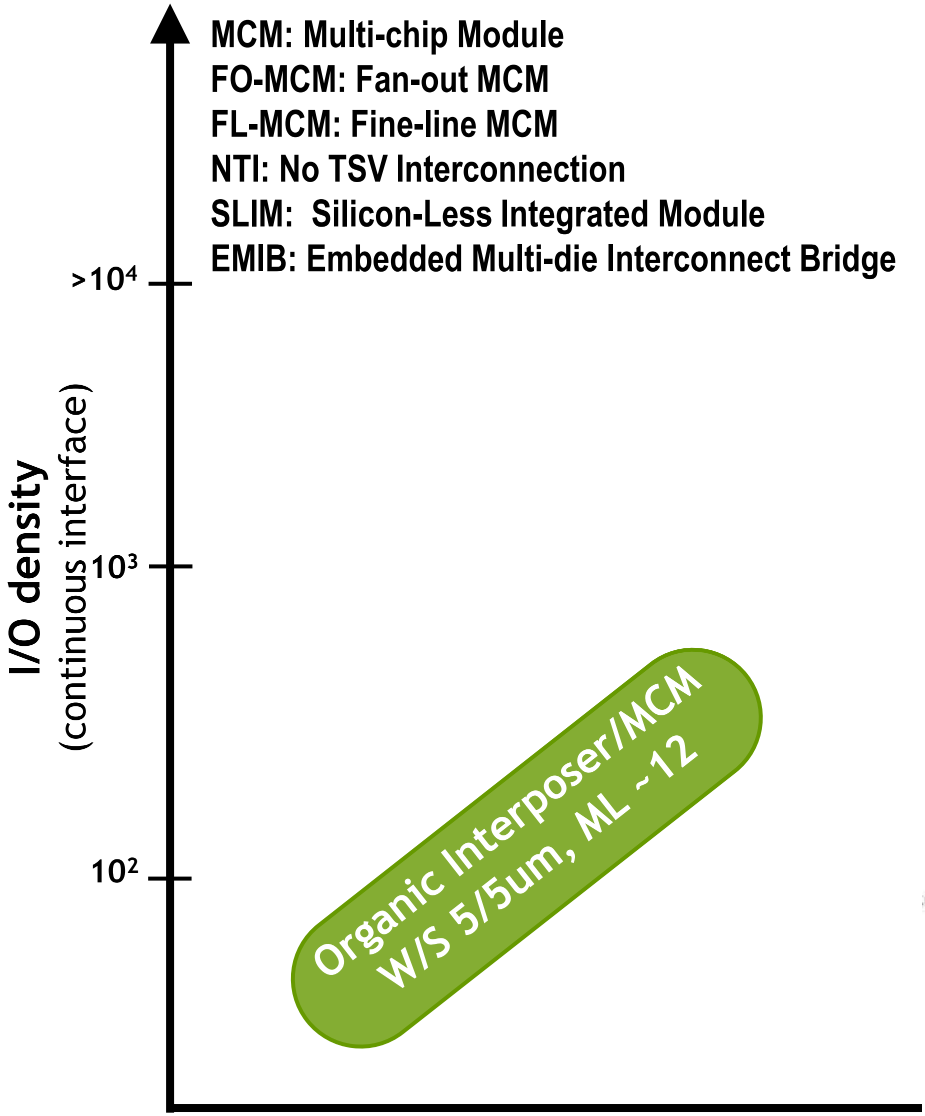
Convey-Xilinx Accelerator





Multi-Die Technology for HBM (Side-by-Side)

- MCM: Multi-chip Module
- FO-MCM: Fan-out MCM
- FL-MCM: Fine-line MCM
- NTI: No TSV Interconnection
- SLIM: Silicon-Less Integrated Module
- EMIB: Embedded Multi-die Interconnect Bridge



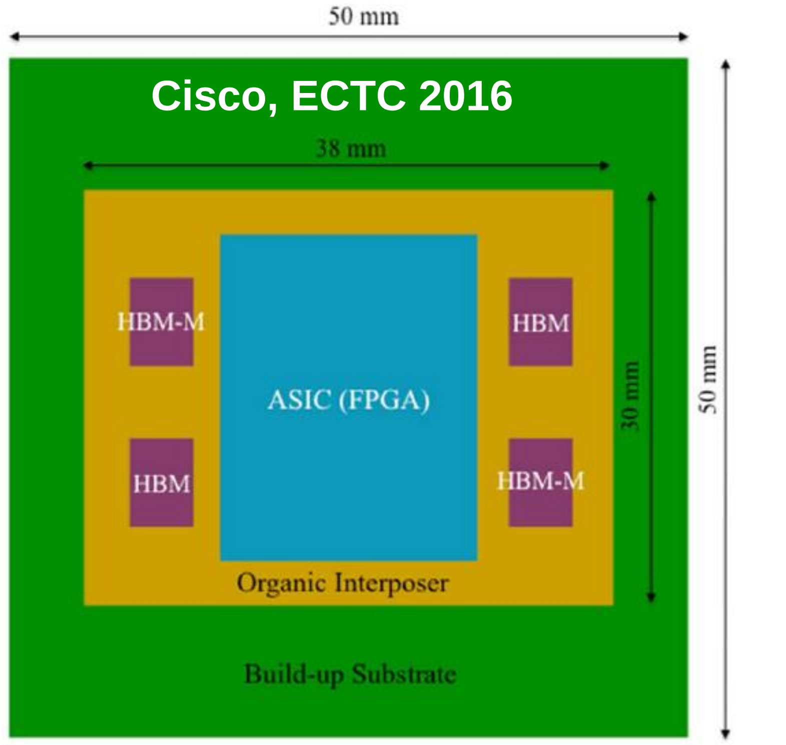

Cisco, ECTC 2016

Figure 2. A schematic top view of the 3D SiP designed.

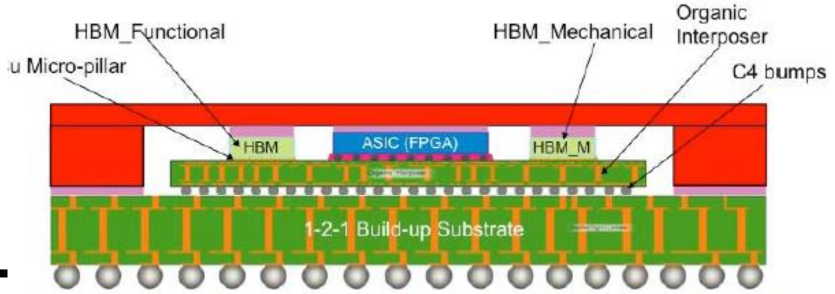
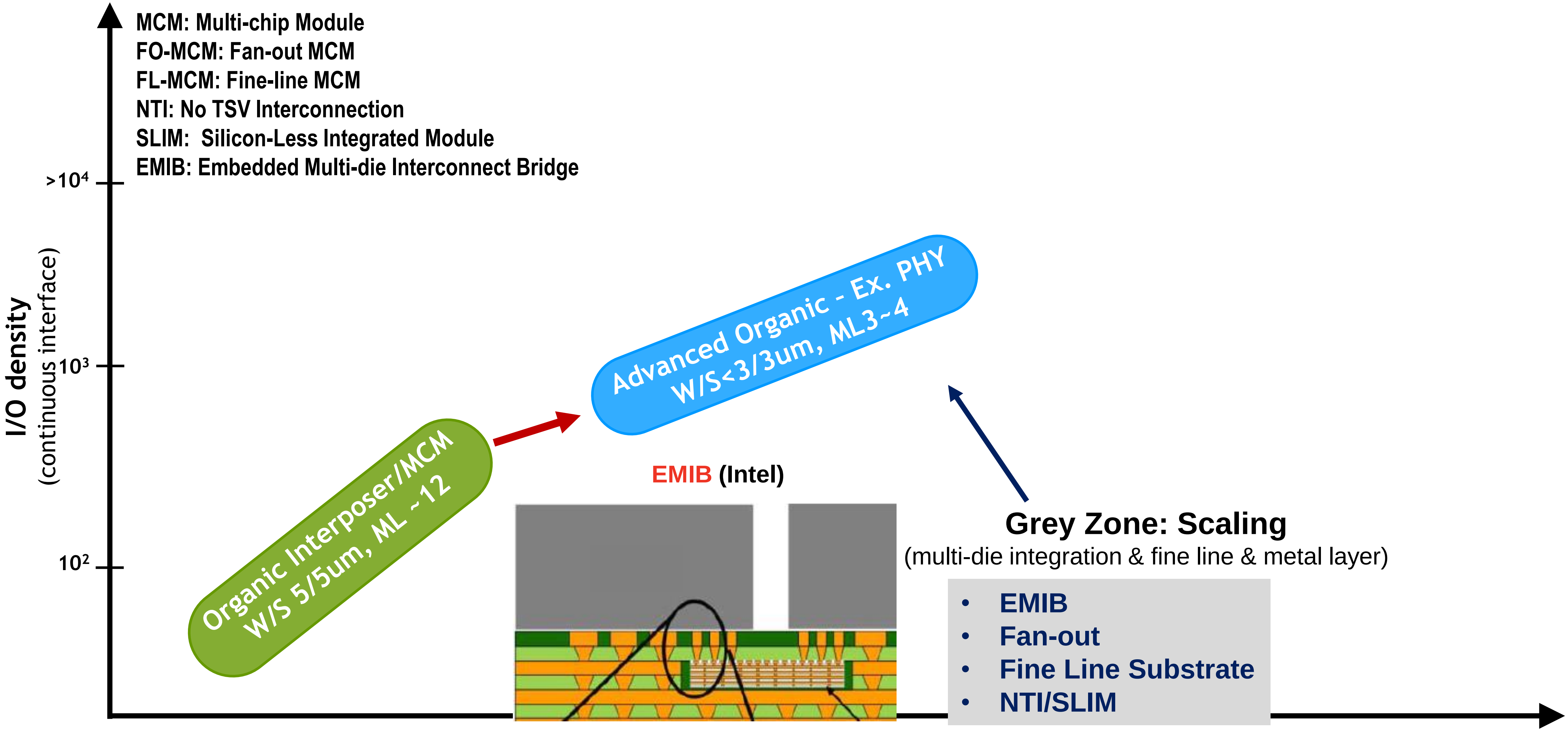


Figure 3. A schematic cross-sectional view of the 3D SiP designed.

Multi-Die Technology for HBM (Side-by-Side)



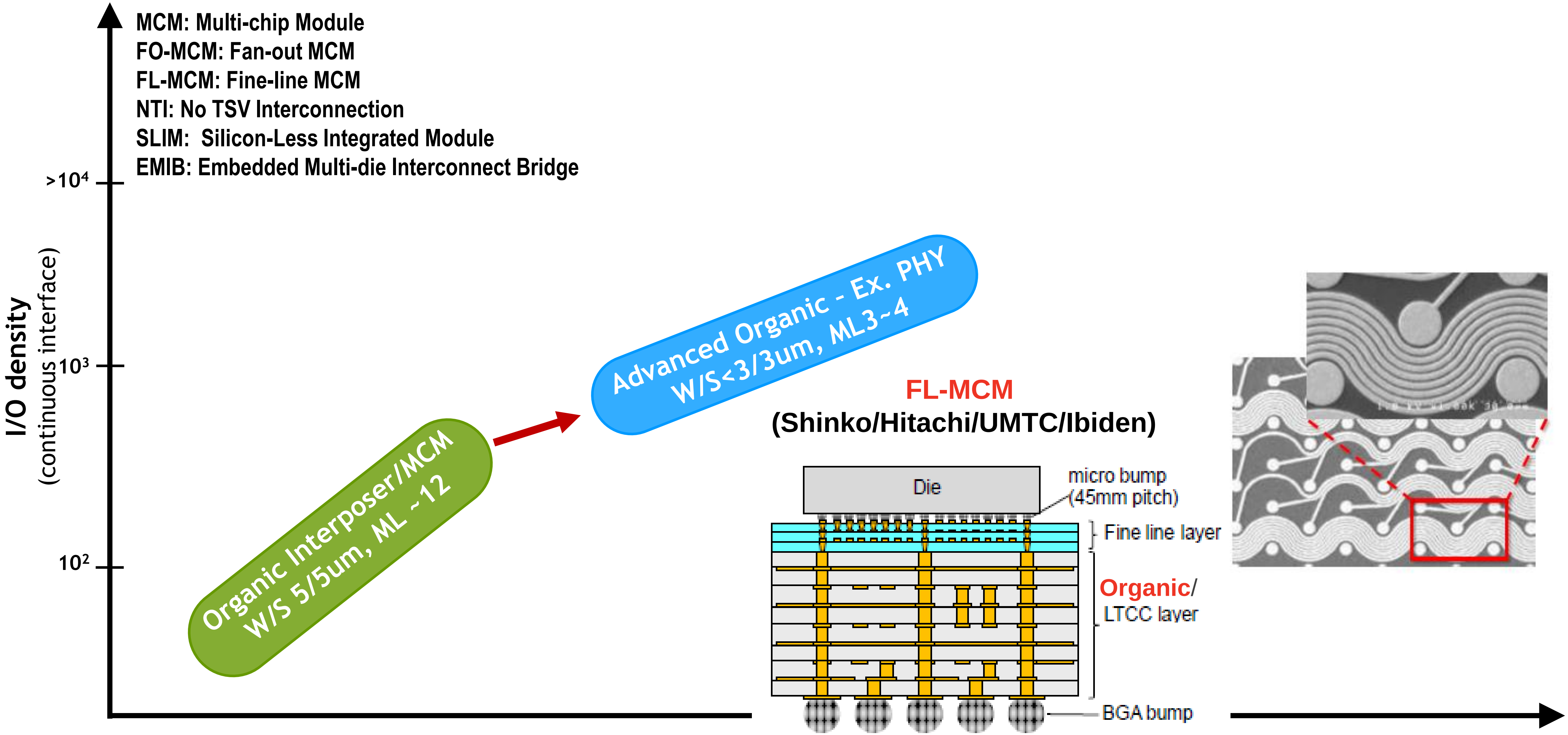
MCM: Multi-chip Module
FO-MCM: Fan-out MCM
FL-MCM: Fine-line MCM
NTI: No TSV Interconnection
SLIM: Silicon-Less Integrated Module
EMIB: Embedded Multi-die Interconnect Bridge

Grey Zone: Scaling
(multi-die integration & fine line & metal layer)

- EMIB
- Fan-out
- Fine Line Substrate
- NTI/SLIM

Multi-Die Technology for HBM (Side-by-Side)

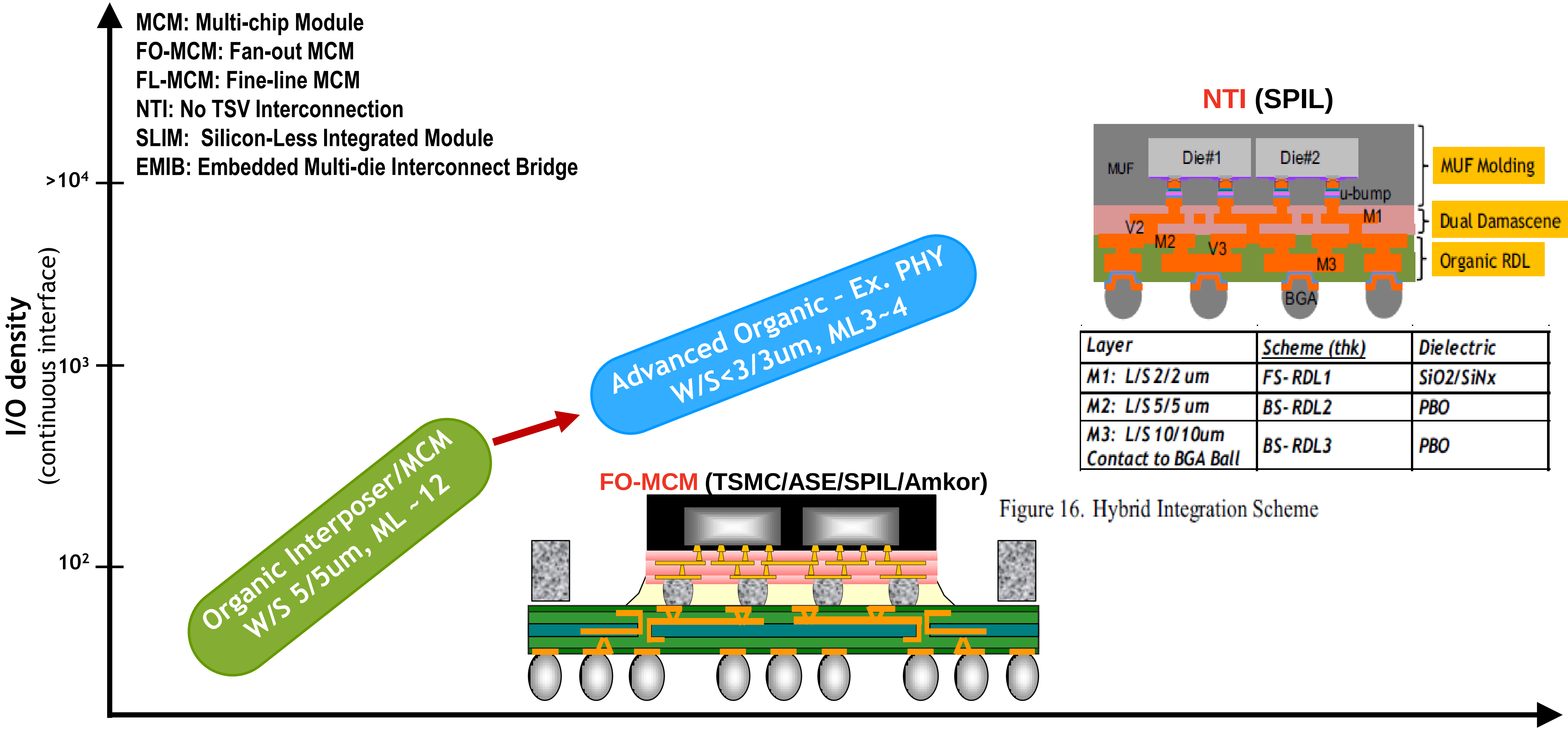
MCM: Multi-chip Module
FO-MCM: Fan-out MCM
FL-MCM: Fine-line MCM
NTI: No TSV Interconnection
SLIM: Silicon-Less Integrated Module
EMIB: Embedded Multi-die Interconnect Bridge



© Copyright 2016 Xilinx

Multi-Die Technology for HBM (Side-by-Side)

MCM: Multi-chip Module
FO-MCM: Fan-out MCM
FL-MCM: Fine-line MCM
NTI: No TSV Interconnection
SLIM: Silicon-Less Integrated Module
EMIB: Embedded Multi-die Interconnect Bridge

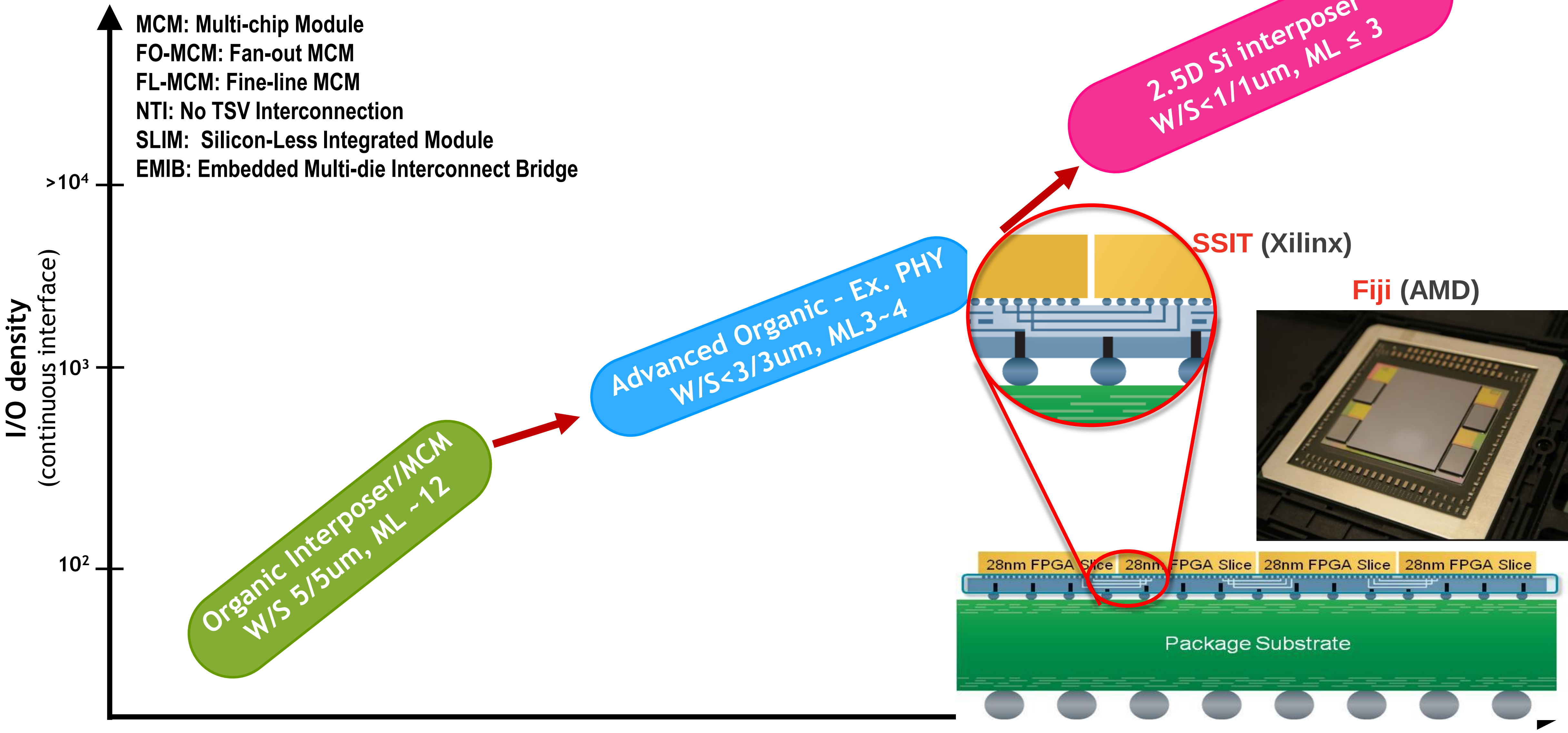


Layer	Scheme (thk)	Dielectric
M1: L/S 2/2 um	FS-RDL1	SiO2/SiNx
M2: L/S 5/5 um	BS-RDL2	PBO
M3: L/S 10/10um Contact to BGA Ball	BS-RDL3	PBO

Figure 16. Hybrid Integration Scheme

Multi-Die Technology for HBM (Side-by-Side)

MCM: Multi-chip Module
FO-MCM: Fan-out MCM
FL-MCM: Fine-line MCM
NTI: No TSV Interconnection
SLIM: Silicon-Less Integrated Module
EMIB: Embedded Multi-die Interconnect Bridge



Organic Interposer/MCM
W/S 5/5um, ML ~12

Advanced Organic - Ex. PHY
W/S<3/3um, ML3~4

2.5D Si interposer
W/S<1/1um, ML ≤ 3

SSIT (Xilinx)

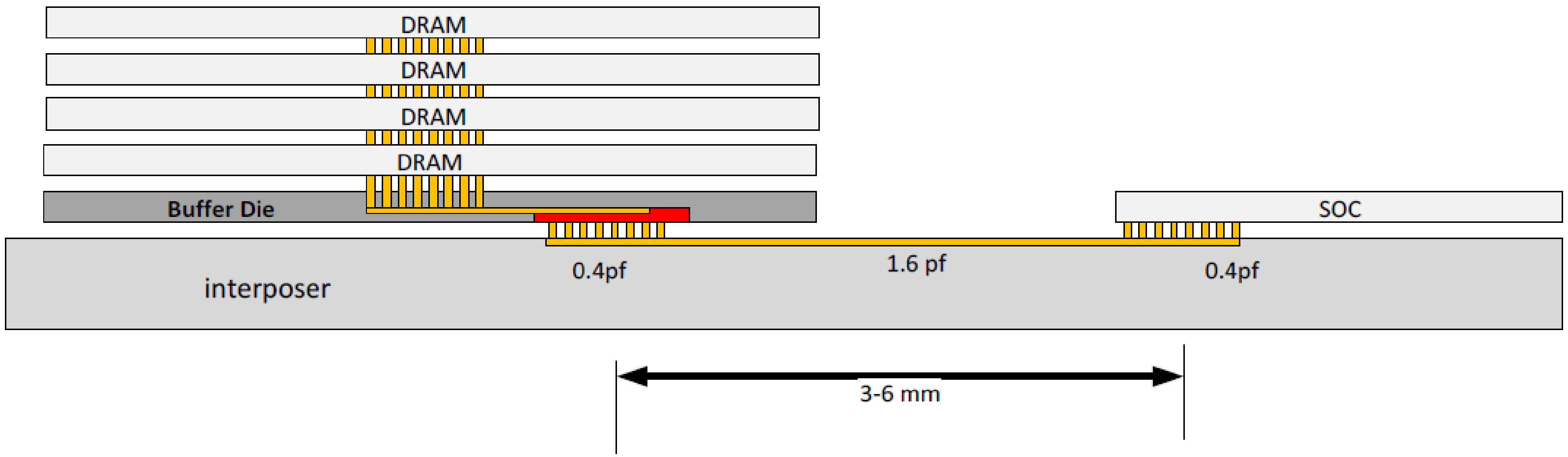
Fiji (AMD)

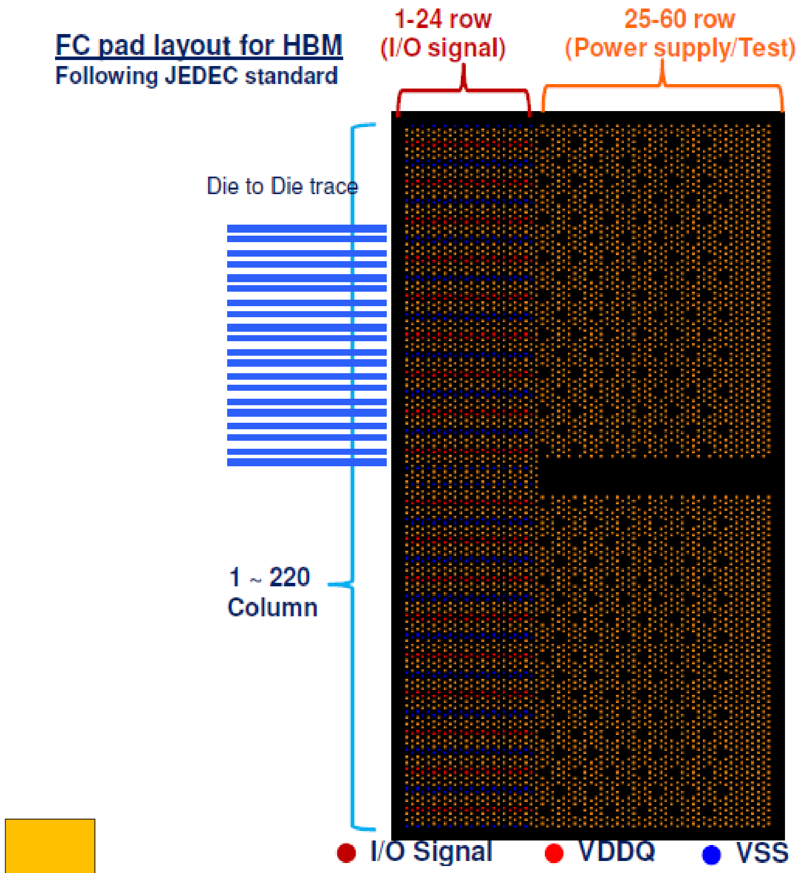
Multi-Die Package Design Rule Comparison

Design Rules for Die to Die interconnection	MCM (Substrate) Integrated Fine Layers	EMIB (Embedded Multi-die Interconnect bridge)	Silicon Interposer (65 nm BEOL)	WLFO (Wafer Level Fan-out)
Minimum Bump pitch (um)	130 (C4) 40 (u-bump d2d interface)	130 (C4) 40 (u-bump) bridge	< 40 (u-bump)	40 um RDL pad pitch
Via size / pad size (um)	10 / 25	0.4 / 0.7	0.4 / 0.7	10/25
Minimum Line & Space (um)	2 / 2	0.4 / 0.4	0.4 / 0.4	2 / 2
Metal thickness (um)	2-5	1	1	2-5
Dielectric thickness (um)	~5	1	1	< 5
# of die-to-die connections per layer + GND shield layer (2L)	1000's	1000's (bridge interface length limited)	10,000's	1000's
Minimum die to die spacing (um)	< 500	<2500	<100	< 250
# of High density layers feasible	Not a limitation 1-3L	Not a limitation	Not a limitation	1-3L layers
Die Sizes for assembly and # of assemblies	Not a concern d2d interconnect only	Size & # limitation?	Not a concern	Size limitation?
In Production	No **(2018)**	No **(2017)**	Yes	No not for 2/2um L/S **(2018)**

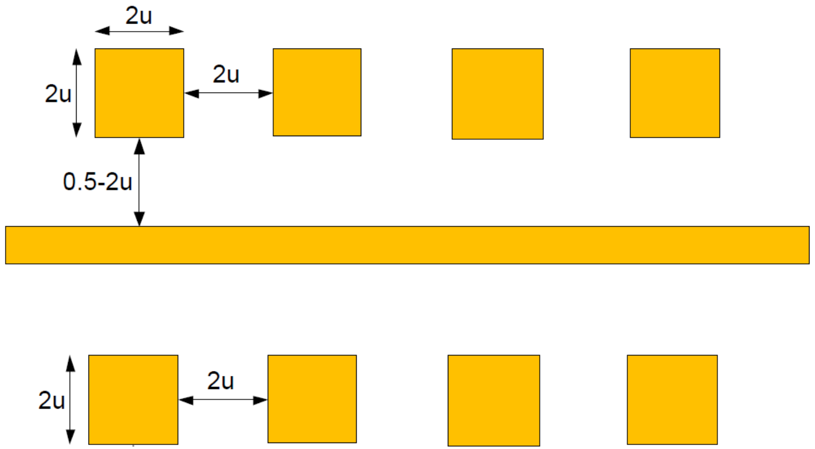
HBM2 System Overview (Jedec)

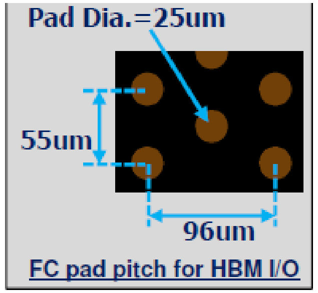
- HBM2 system with SOC/DRAM on interposer with 3-6mm length
- 24 signals across 55um u-bump pitch across interface
- Supports 2Gb/s PHY (1Tb/sec bandwidth for 4-Hi)





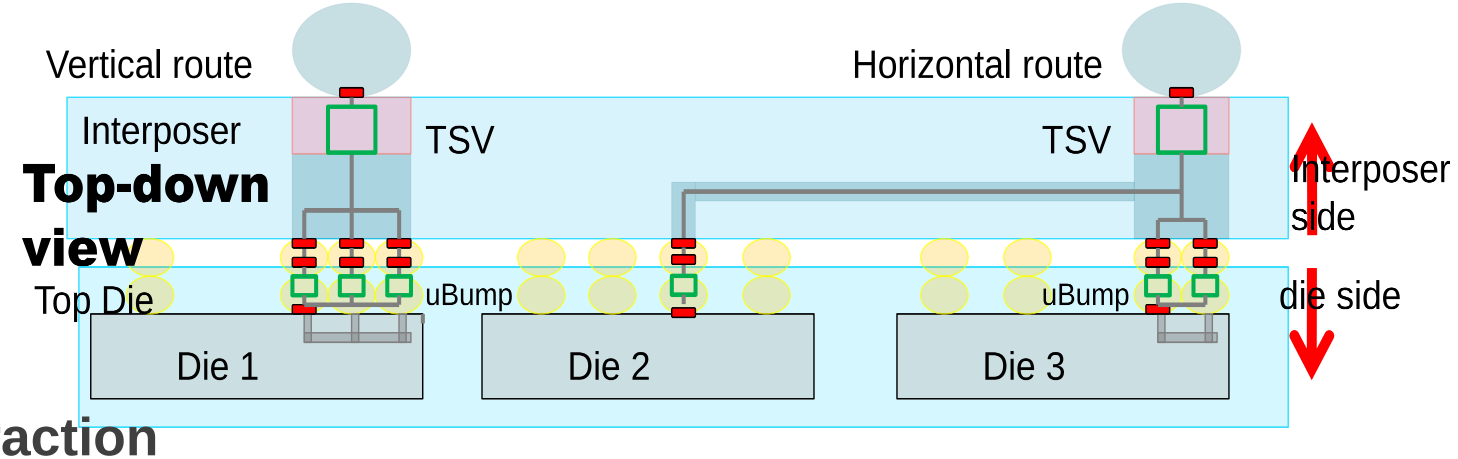
(Jedec) Interposer Parameters *				
Width	Space	Thickness	Length	Resistance
2u	2u	0.5-2u	6mm	36 ohms





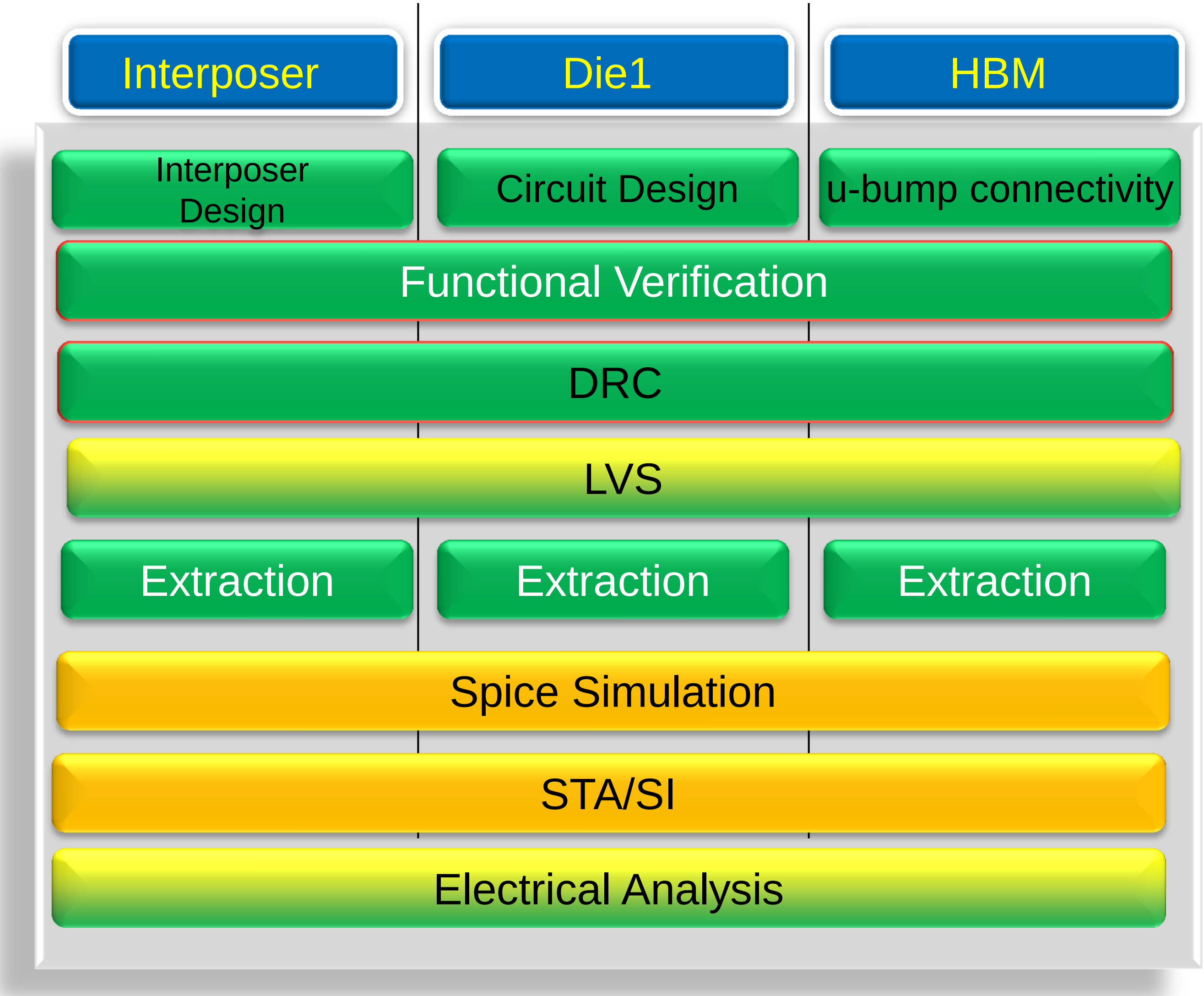
Interposer Design Tools & Methodology

- **Vertical Routing**
  - a model from ubump to package pin is generated and used by high frequency designs (e.g. GT and IO)
- **Horizontal Routing**
  - **Die LVS and extraction**
    - Standard extraction
    - ubump is extracted as a subcircuit
  - **Interposer with box die LVS and extraction**
    - Interposer metal extraction
    - TSV is extracted as a subcircuit
- **Combine the extracted netlists from die and interposer for simulation**



Interposer Design Tools & Methodology

- 2.5D - uses the same tool sets as single die design with customized interposer / top die PDK
- EDA vendor Tools are validated by TSMC design reference flow
- PKG - uses same tool sets as Flip chip (C4-to-BGA)
  - TSV budget is handled in the Silicon design environment
  - Layout and PI tools must be capable to handle large data sets



Interposer | Die1 | HBM

Interposer Design | Circuit Design | u-bump connectivity

Functional Verification

DRC

LVS

Extraction | Extraction | Extraction

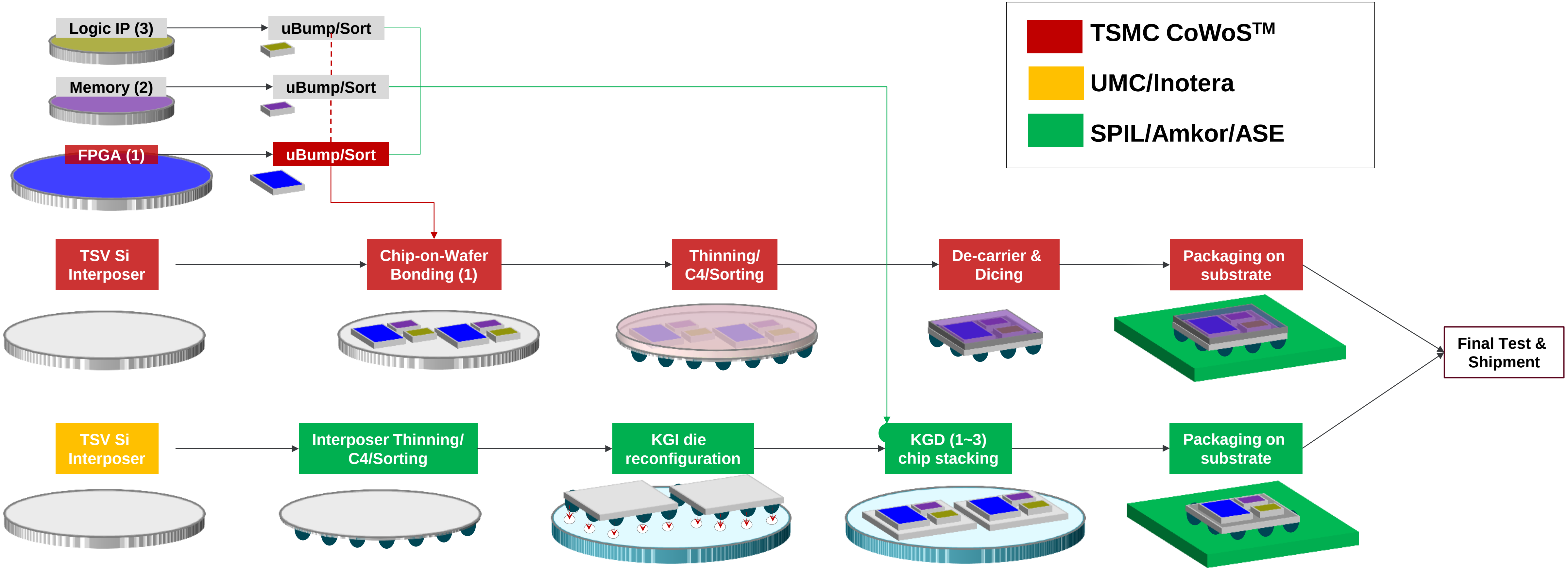
Spice Simulation

STA/SI

Electrical Analysis

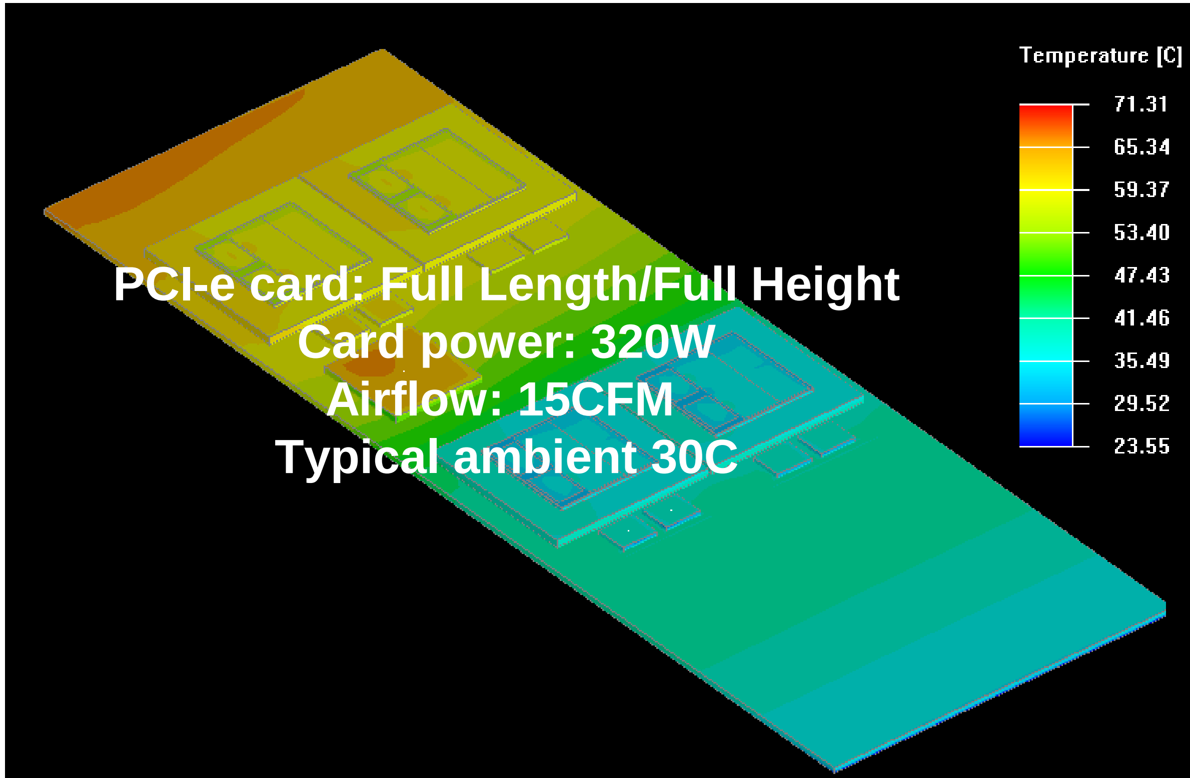
Supply Chain – Silicon Interposer Approach

- Xilinx in production with 2nd generation of products with TSMC CoWoS



*CoC/CoS/CoWoS-last

HBM Integration – HPC Application



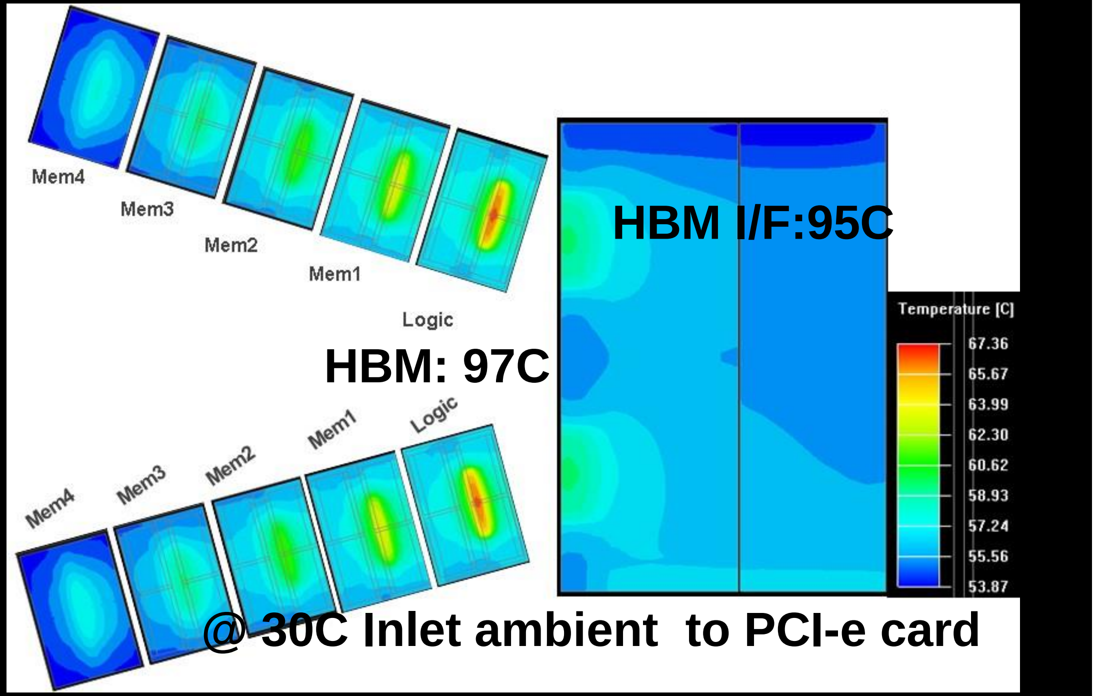
PCI-e card: Full Length/Full Height
Card power: 320W
Airflow: 15CFM
Typical ambient 30C

- HBM Power map provided by vendors
- Thermal model can be done in Flotherm or IcePak environments for example

HBM can be 97C and HBM I/F 96C @30C

HBM gradient ~14C (~2.5C/Layer)

Air cooling can be a challenge! HBM 8-Hi needs to support > 95C T_j …..



HBM I/F:95C

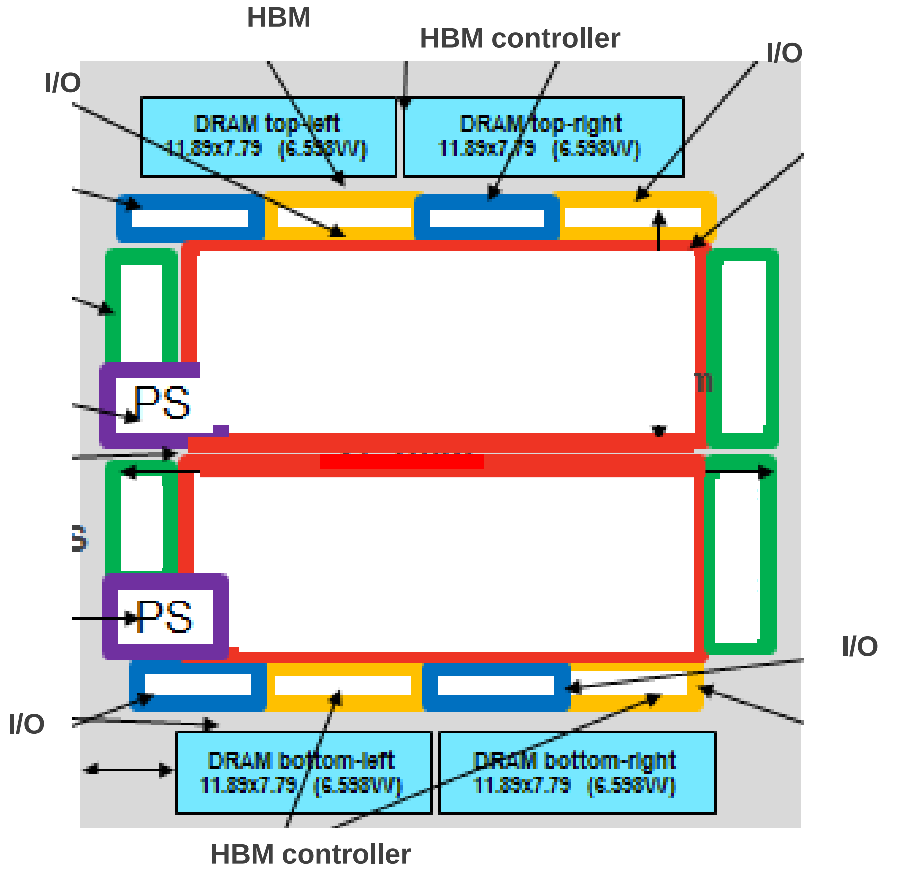
HBM: 97C

@ 30C Inlet ambient to PCI-e card

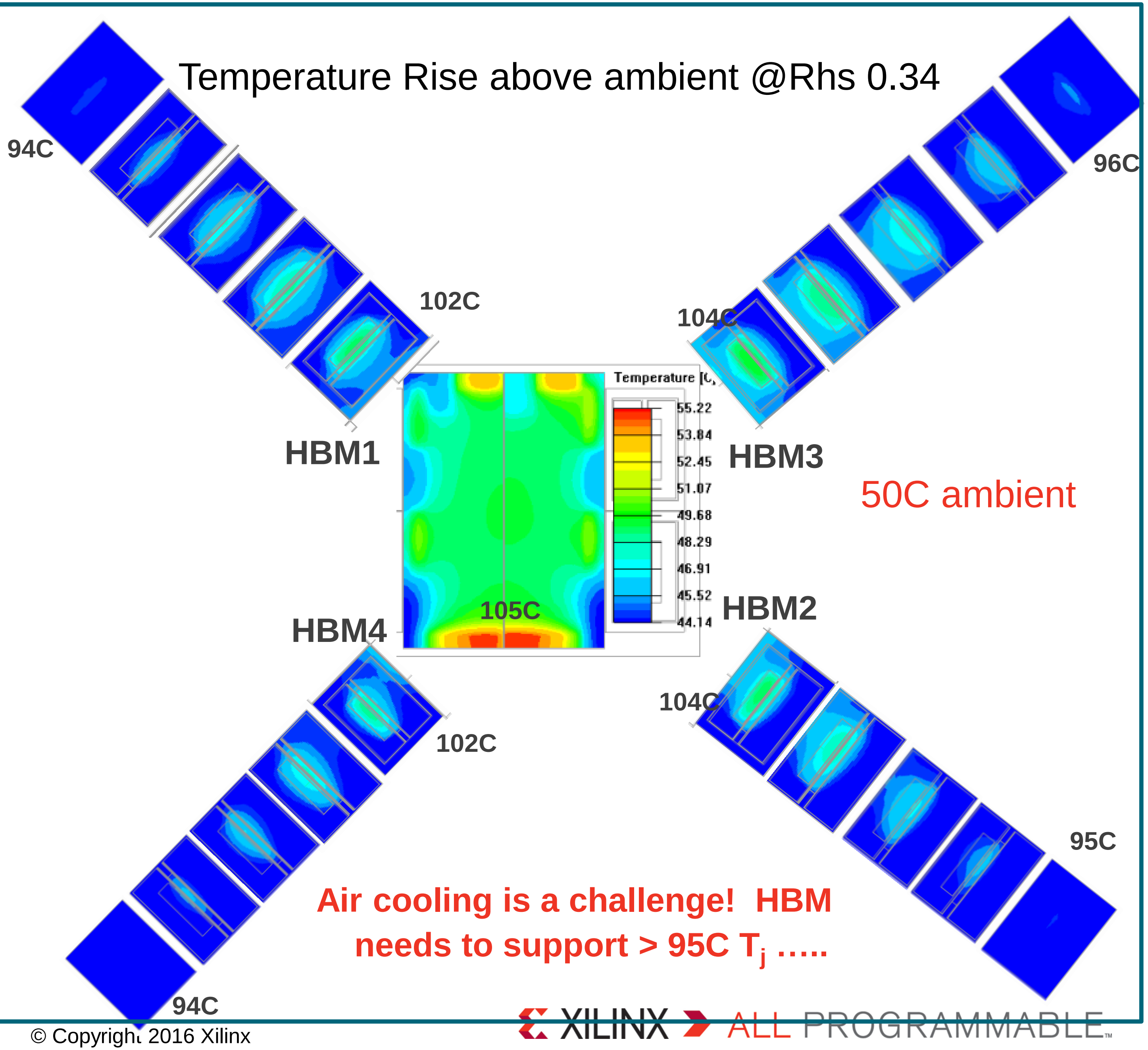
Telecom Application with HBM

PKG size:52.5x52.5 mm

Total Heat Dissipation is 116.3 W



Temperature Rise above ambient @Rhs 0.34



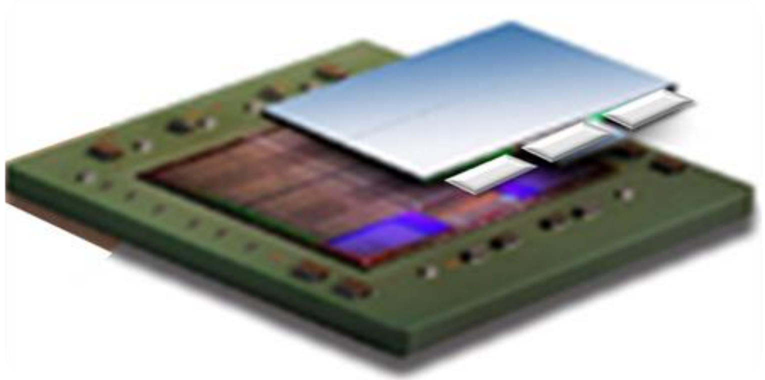
50C ambient

Air cooling is a challenge! HBM needs to support > 95C T_j …..

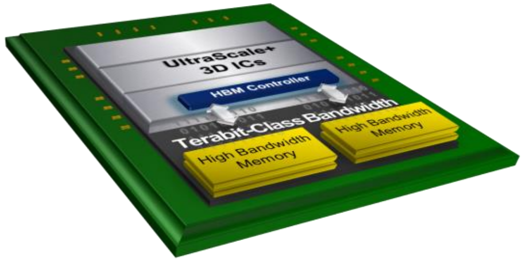
Summary

- Tb/s low latency bandwidth and lower system power is driving the need for HBM adoption

- Silicon Interposer (2.5D) is the incumbent technology of choice. Potentially lower cost, fine pitch interconnect wafer-level and substrate based technologies are emerging



- To drive broader adoption of HBM applications (cooling limited) and higher performance stacks (8-Hi), higher HBM junction temperature (>95C) needs to be supported



Follow Xilinx

XILINX ALL PROGRAMMABLE™

- facebook.com/XilinxInc
- twitter.com/XilinxInc
- youtube.com/XilinxInc
- linkedin.com/company/Xilinx
- plus.google.com/+Xilinx