Coupling Extraction and Optimization for Heterogeneous 2.5D Chiplet-Package Co-Design

MD Arafat Kabir[1], Dusan Petranovic[2], Yarui Peng[1]
University of Arkansas, Fayetteville, AR, US[1]
Mentor, a Siemens Business, Fremont, CA, US[2]

MD Arafat Kabir

MD Arafat Kabir joined in Computer Science and Computer Engineering department at the University of Arkansas in the Fall of 2018 as a Ph.D. student. He received his B.Sc. in electrical and electronic engineering (EEE) from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh in 2017.

His research interests are in Electronic Design Automation (EDA) and Very Large-Scale Integration (VLSI) design. Currently, he is working on an NSF funded project. The project is to develop tool flow for design, extraction, and optimization of multi-chip fan-out wafer-level-packaging for low-power heterogeneous systems.

https://e3da.csce.uark.edu/~makabir/

UNIVERSITY OF ARKANSAS

Introduction

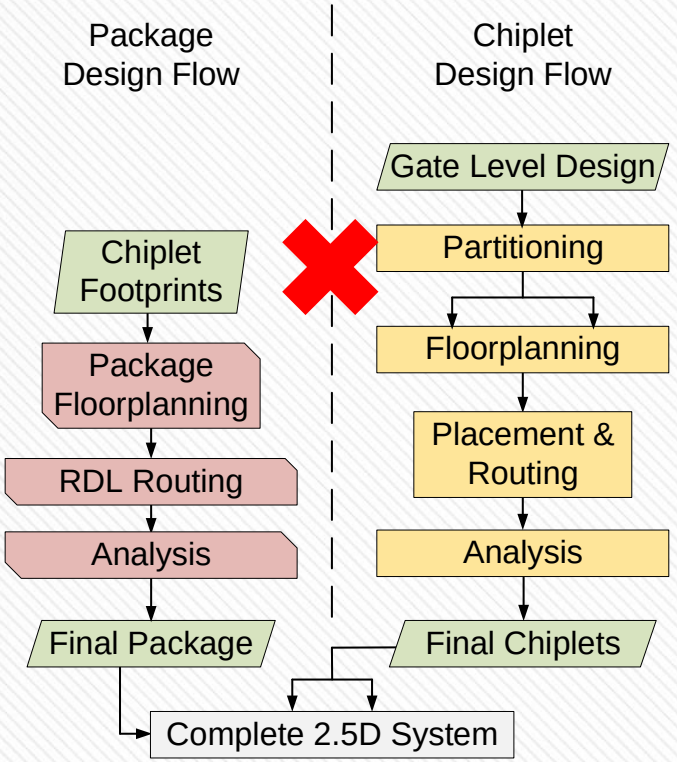
- **Package becomes increasingly critical in post-Moore's Law era**
  - Transistor scaling is saturated, and chips are reaching reticle limit.
  - 2.5D and 3D packages provide high bandwidth and compact size.
- **2.5D packaging opens a new door of opportunities**
  - Novel design techniques like plug-and-play, Drop-in, Hardware security
  - Heterogeneous integration capabilities (AMD EPYC2, Intel Lakefield)
- **Need for a cross-boundary package-aware design strategy**
  - Interactions between the package and chiplets are significant
  - Optimization goals for 2.5D system are different: inductance, signal integrity, thermal, reliability, cost, turnaround time, flexibility, etc.
- **Objectives**
  - Study of holistic iterative optimization flow for homogenous systems
  - Study and comparison of in-context flow for heterogenous systems

UNIVERSITY OF ARKANSAS

Need for Holistic Consideration

❑ Need for a cross-boundary package-aware design strategy:

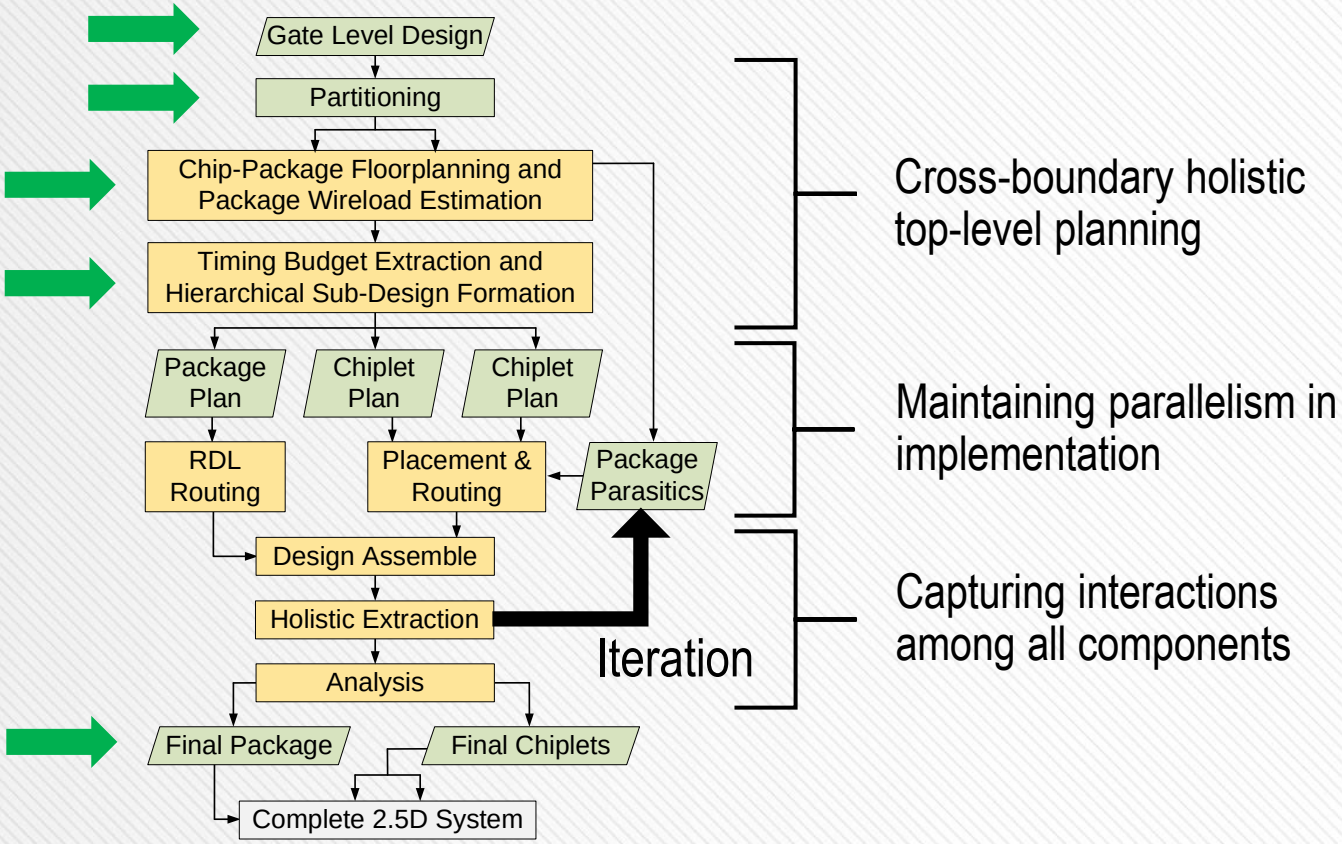
- Traditional die-by-die flow treats each chiplet as an individual design.
- Partition tool needs to be aware of the delay introduced by redistribution layers (RDLs) with detailed parasitic extraction
- Package planning tool may need to modify chiplet pin arrangements to optimize RDL routing and package floorplan.
- Holistic chiplet-package interactions can be used to optimize and fine tune system performance and ensure reliability.



Traditional Flow

Our Holistic Flow

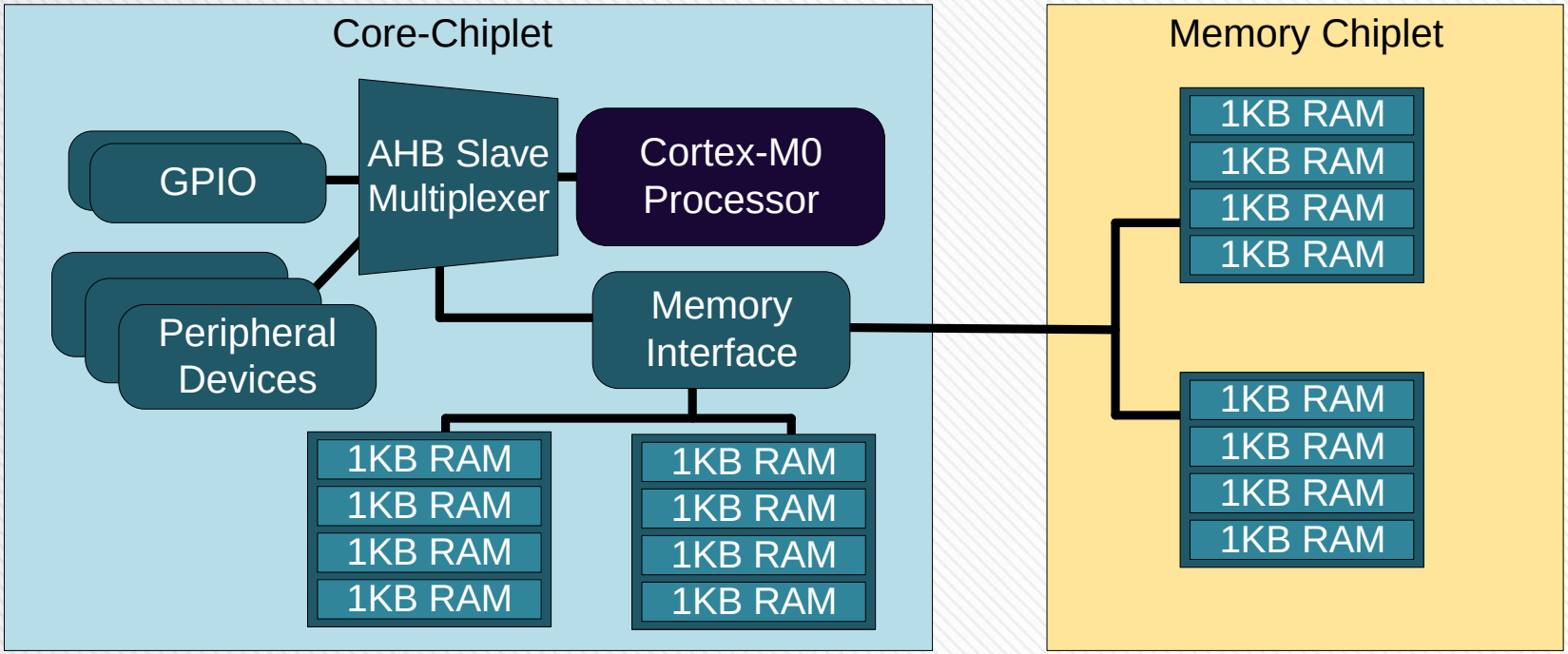
- **Exchange of cross-boundary design information in planning, design, analysis, and optimization steps**



Our Holistic Flow

MCU Architecture and Partitions

- **System architecture of our test design**
  - Microcontroller system based on ARM Cortex-M0 core
  - 16KB RAM with some common peripheral devices

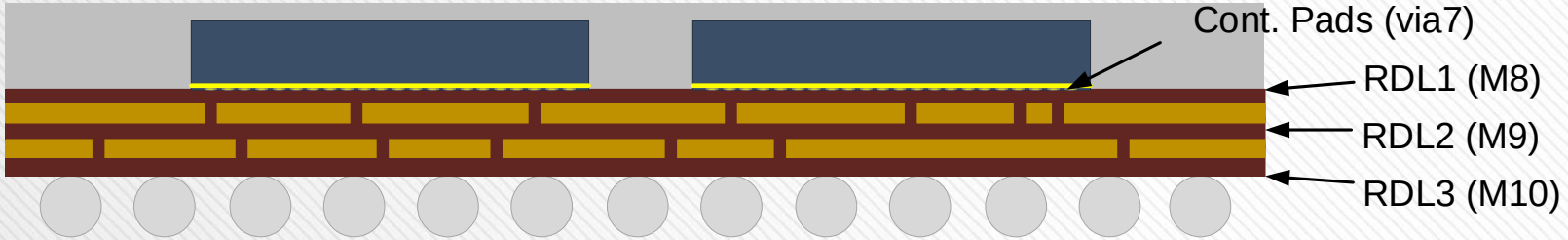


System architecture and chiplet partitions

Technology Settings

- **We use Nangate45nm as our PDK**
  - M1-M7 used for chiplet routing
- **We modify the top three layers to include 2.5D package RDLs**
  - Dimensions are similar to the TSMC 2.5D InFO technology

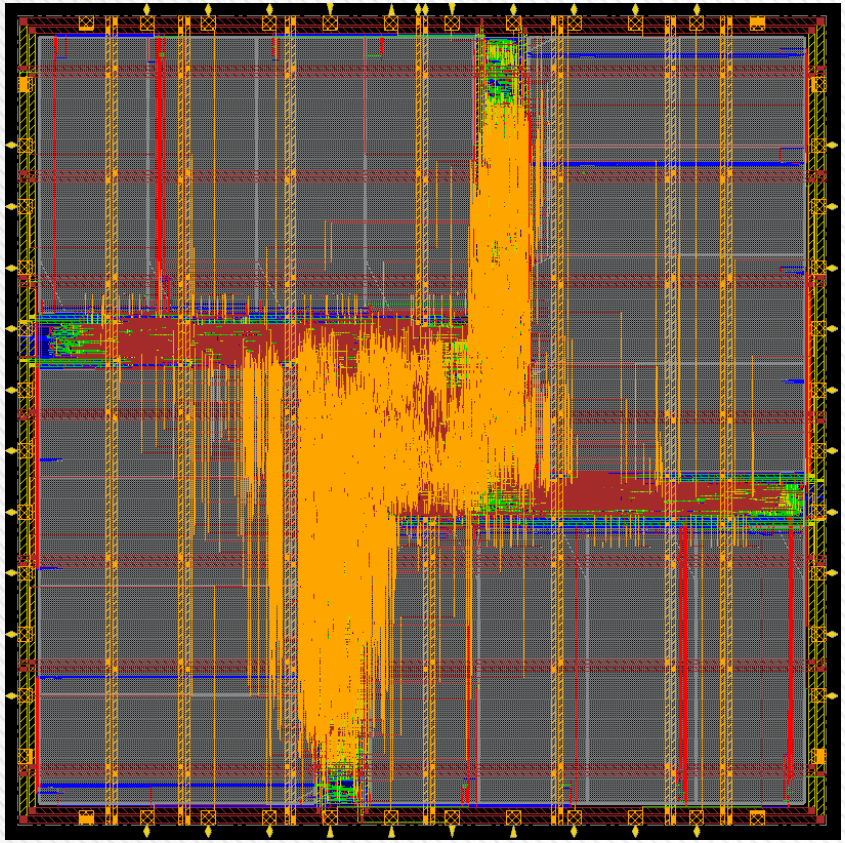
	M6	via6	M7	via7	RDL1	viar1	RDL2	viar2	RDL3
Height	2.28	3.08	3.9	7.5	12.5	17.5	22.5	27.5	32.5
Thickness	0.8	0.82	3.6	5	5	5	5	5	5
Width	0.4	0.4	2	5	10	10	10	10	10
Spacing	0.4	0.44	2	10	10	20	10	20	10



UNIVERSITY OF ARKANSAS

Reference 2D Design

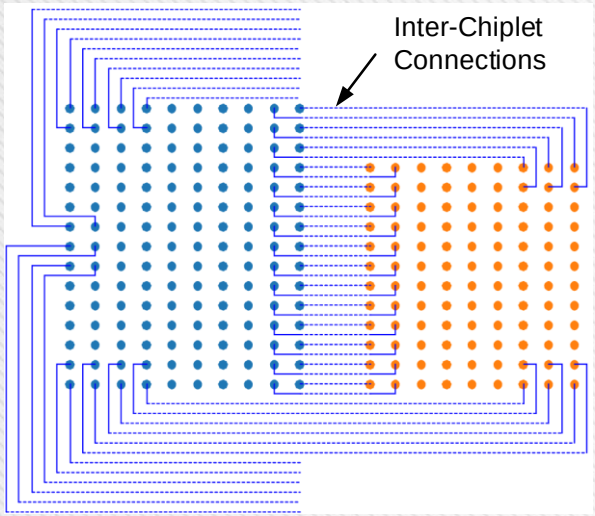
- For comparative study, we implement the microcontroller as a 2D system using conventional 2D chip design flow
  - Only with 16KB memory flavor



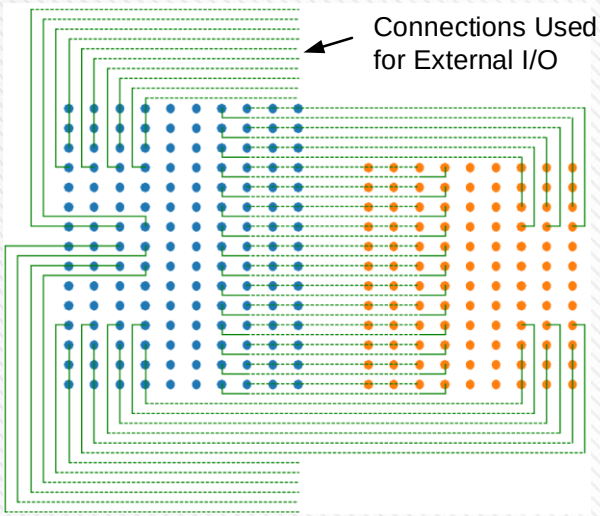
UNIVERSITY OF ARKANSAS

Co-Planning Strategy

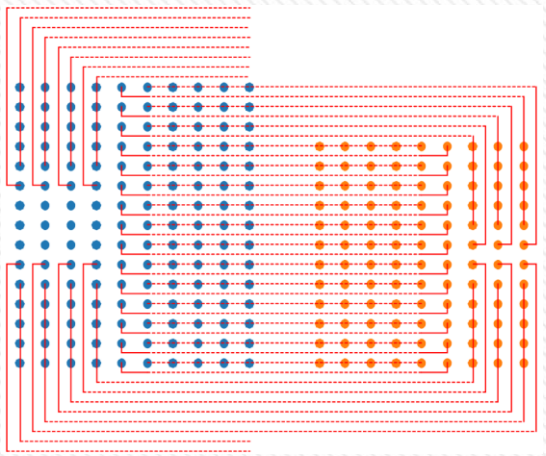
- **In-house RDL planning tool implements our planning strategy.**
- **Our planning strategy,**
  - Package floorplan and routing are determined before signal assignment
  - Use of short and straight wires to route the chiplet pins
  - Greedy signal assignments of pins based on package wire-length and timing report from synthesis tools



RDL1



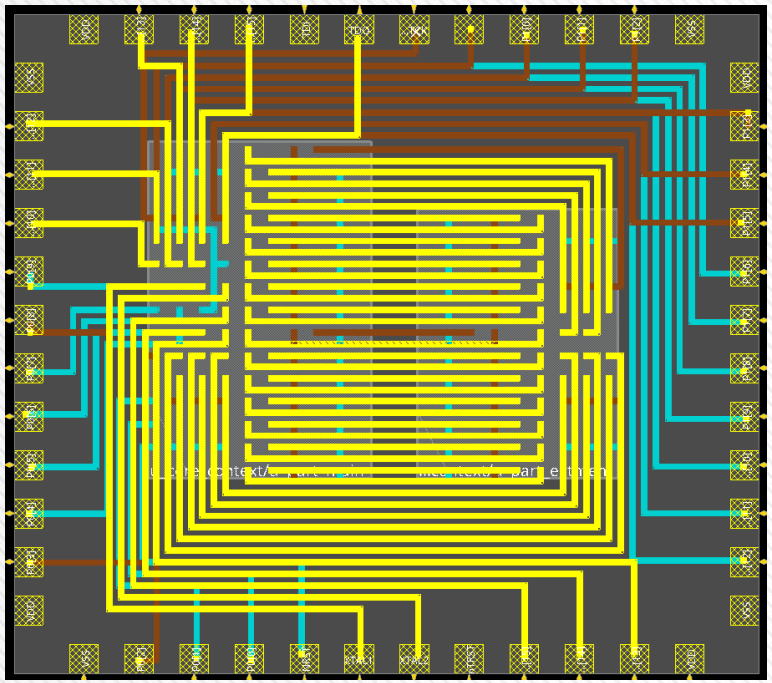
RDL2



RDL3

Top-Level Plan and Package Design

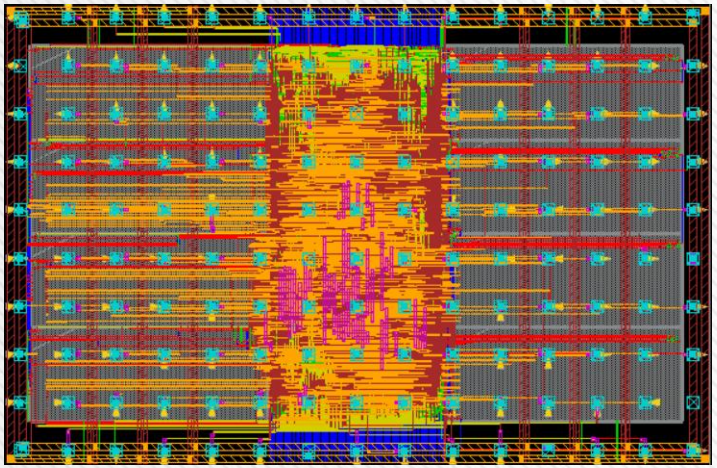
- **RDL planner tool generates routing scripts**
  - Package routing is completed using routing scripts
  - Timing budgets are extracted, and hierarchical sub-designs are formed
  - The estimated package wireload is used in sub-design constraint files



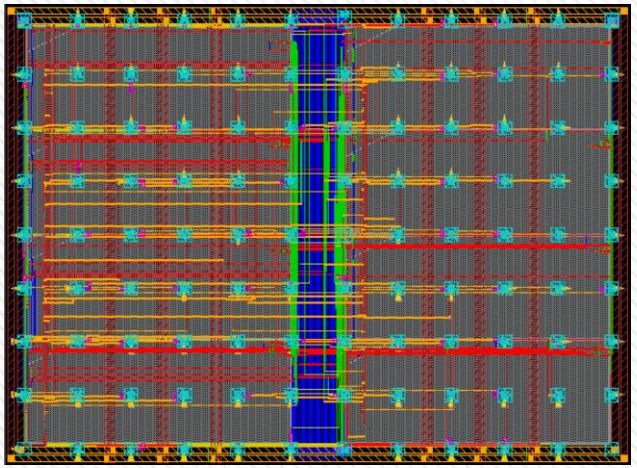
UNIVERSITY OF ARKANSAS

Placement & Routing

- **After top level planning, chiplets and package are implemented independently with constraints propagated from top-level**
  - Top level design is divided into hierarchical sub-designs.
  - Chiplet floorplan and PDN are finalized.
  - Chiplet implementation is the same as conventional 2D chip that includes power planning, placement, time design, routing and post routing optimizations.



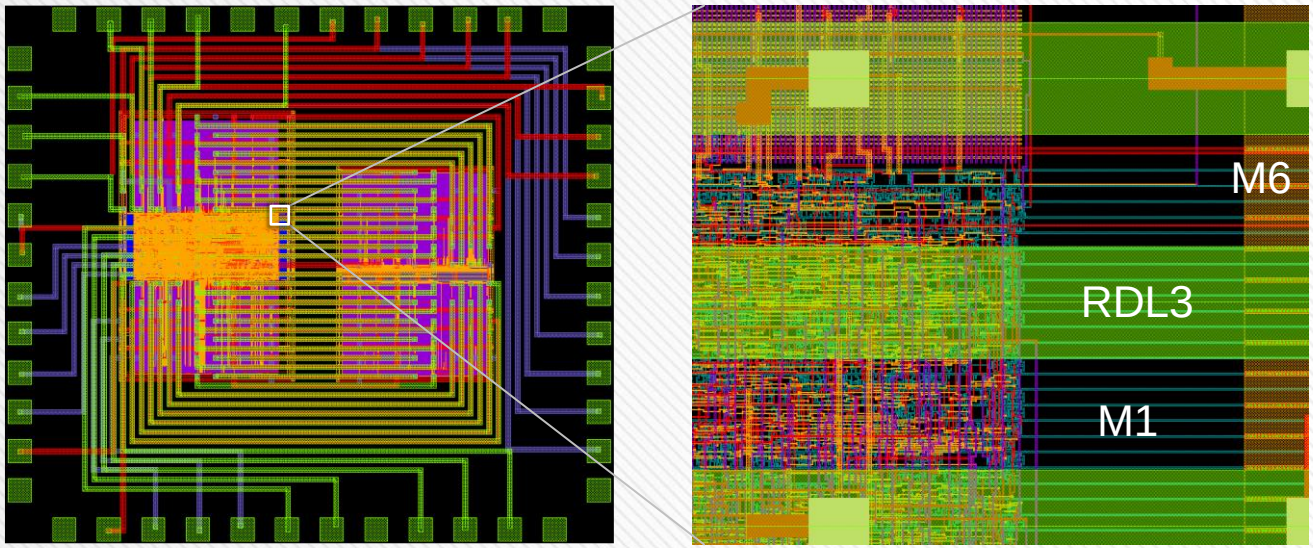
Core-Chiplet



Mem-Chiplet

Design Assembly & Holistic Extraction

- **Finished package and chiplet designs are assembled for holistic extraction.**
  - As the design environment has everything together, incremental optimizations can be performed to improve overall system performance.
  - The analysis and optimization tools have all the information needed to account for the impacts of RDLs on chiplet design.



Assembled System

Zoom-in View

Holistic Extraction Result

Chiplet-Package coupling capacitance

- The columns for RDL1, RDL2, and RDL3 show the coupling capacitance between package layers and chiplet layers (in fF).
- The coupling of package layers with M7 is low because of a smaller number of wires on M7. However, there exists significant coupling with the wires on M6, which is captured in the parasitic extraction process
- Also, M7 and RDL1 are extracted with considerations from the other side

Coupling Capacitance (CCAP)						
Metal Layer	**M1-M5**	**M6**	**M7**	**RDL1**	**RDL2**	**RDL3**
M1-M5	6120	442.2	28.65	52.95	8.102	5.862
M6	442.2	596.6	78.03	122.8	12.98	10.53
M7	28.65	78.03	30.63	15.02	1.509	2.256
RDL1	52.95	122.8	15.02	299.3	1016	39.06
RDL2	8.102	12.98	1.509	1015	298.3	1085
RDL3	5.862	10.53	2.256	39.06	1084	578.4
Ground Capacitance (GCAP)						
Metal Layer	M1-M5	M6	M7	RDL1	RDL2	RDL3
Capacitance	21119	2054	272	1040	247	636

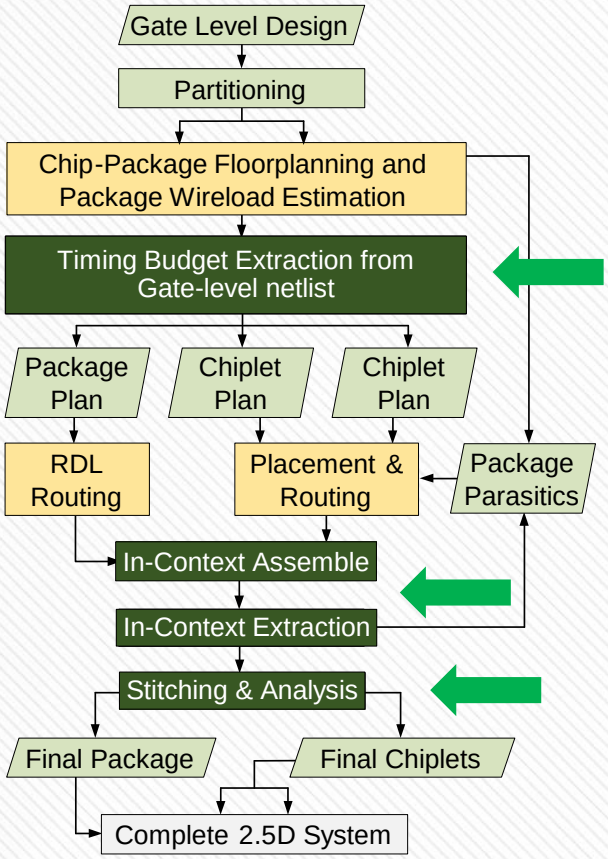
Iterative Optimization Results

- **Chiplet-package interaction is used to improve the system performance through iterative optimizations**
  - Chiplet design tool automatically optimizes the inter-chiplet IO buffers to compensate for package overhead by 62.5%

Design Case	Chiplet Design	Logic Gates#	Buffer/ Inverter#	Die Size (um^2)	M6 WL (mm)	M7 WL (mm)	Power (mW)	Freq. (MHz)	Freq. Overhead
Case-1	2D Chip	17595	3700	550x550	79.94	0	10.6	333	**0%**
Case-2	Core	17783	2740	390x590	30.81	1.783	7.751	245	**100%**
	Mem	132	132	350x470	5.986	0.598	0.194		
Case-3 initial	Core	17915	2865	390x590	31.86	1.875	9.043	280	**60.23%**
	Mem	148	148	350x470	8.201	0.589	0.216		
Case-3 final	Core	18214	2955	390x590	31.42	2.02	9.840	300	**37.50%**
	Mem	45	45	350x470	8.445	0.624	0.162		

In-Context for Heterogeneous

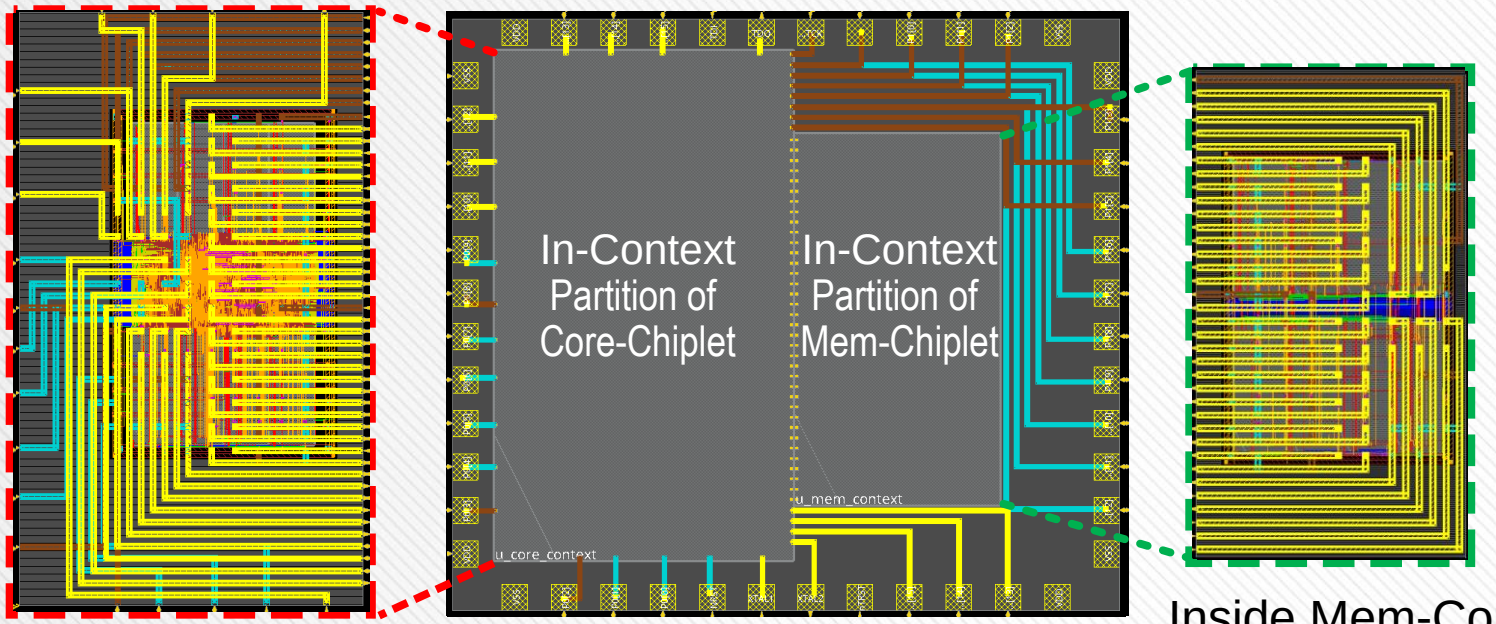
- **Existing EDA tools cannot handle multiple heterogeneous technologies together in a common design scope**
  - Holistic timing budget and parasitic extraction not possible for heterogeneous technologies.
- **In-Context Design Flow**
  - Break down the package into package contexts and create an extended partition for each chiplet.
  - Perform in-context extraction and then stitch all SPEFs in the analysis tool for accurate timing context and analysis.
  - Cross-boundary in hierarchy levels



Our In-Context Flow

In-Context Partitions

- **An extra level in the design hierarchy is created combining each chiplet and its surrounding package regions**
  - Heterogeneous technologies would require a top-level design for each technology to create these in-context partitions.



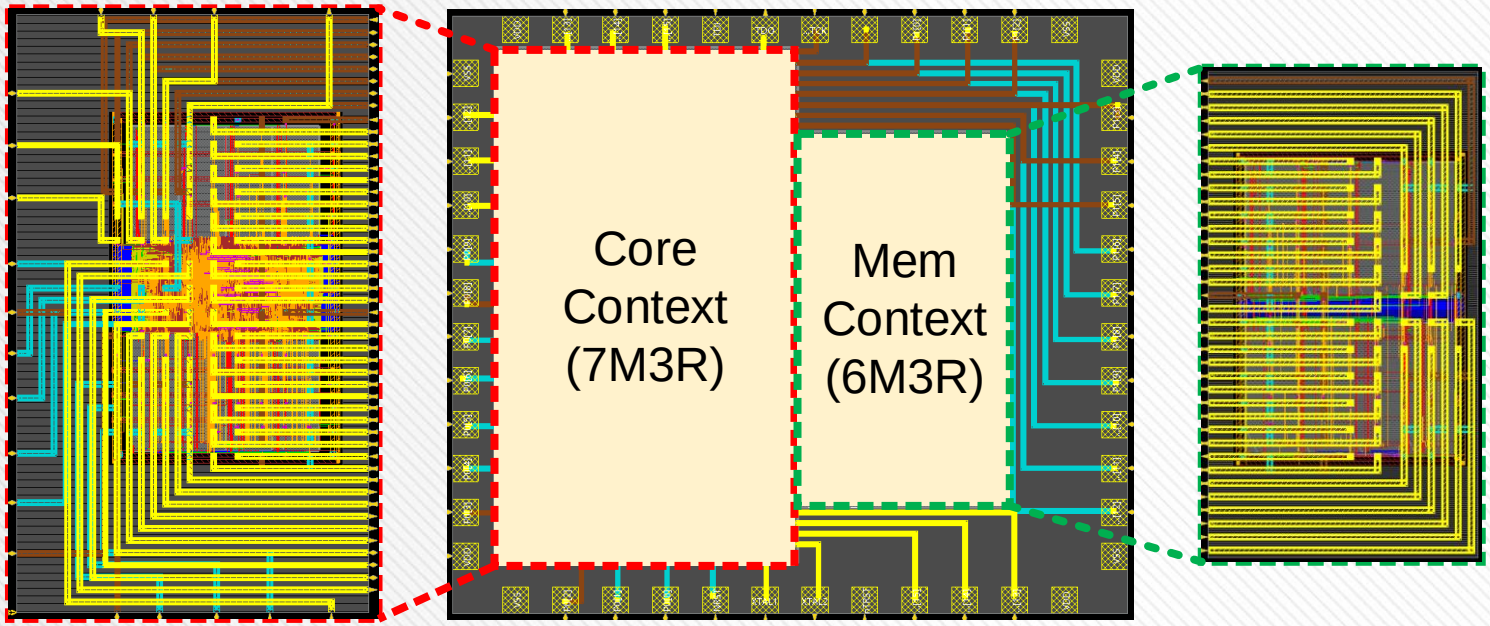
In-Context Extraction Comparison

- **We perform in-context extraction on the homogeneous design for comparative study**
  - The total GCAP error is only 0.71% and total CCAP error is only 0.79%
  - InC package GCAP is overestimated due to fringe cap on cutting edges
  - Die-by-die extraction (DbD) overestimates GCAP and underestimates CCAP on all layers which may cause signal integrity issues

	Metal Layer	M1-M5	M6	M7	R1	R2	R3
GCAP	Holi	21119	2054	272	1040	247	636
	DbD	21139	2090	278	1539	362	658
	InC	21119	2053	273	1103	306	696
	DbD Err	0.10%	1.78%	2.09%	47.97%	46.77%	3.45%
	InC Err	0.00%	-0.01%	0.09%	6.03%	24.0%	9.46%
CCAP	Holi	9172	1263	156	1544	2421	1721
	DbD	9125	1213	141	1378	2287	1699
	InC	9171	1265	153	1563	2489	1765
	DbD Err	-0.52%	-3.95%	-9.94%	-10.75%	-5.55%	-1.30%
	InC Err	-0.01%	0.17%	-2.10%	1.20%	2.81%	2.56%

A Heterogeneous Design

- **We prepare a proof-of-concept heterogeneous design using various metal stacks and cell libraries for chiplets**
  - Package is routed using three RDLs (3R)
  - Core-chiplet uses seven chip-routing layers (7M) and Nangate Cell library
  - Memory-chiplet uses six chip-routing layers (6M) and GSCL Cell library



UNIVERSITY OF ARKANSAS

In-Context Iterative Optimization

- **Using in-context flow, we perform design and iterative optimizations of the chiplets**
  - The initial iteration with estimated wireload achieved same performance as that of the initial iteration of holistic design.
  - Final iteration achieved 298 MHz maximum frequency which is almost as good as 300 MHz of the holistic design.

Design Case	LPD	Max Frequency	Total Power
With RDL wireload	3.55 ns	281 MHz	10.5 mW
In-Context 1st iteration	3.35 ns	298 MHz	10.6 mW
In-Context 2nd/final	3.35 ns	298 MHz	10.6 mW
Holistic final design	3.33 ns	300 MHz	10.7 mW

UNIVERSITY OF ARKANSAS

Conclusions & Future Work

Conclusions

- Chiplet-Package interactions need to be considered in 2.5D systems
- Our flow effectively captures the impact of RDLs in optimization and analysis steps.
- It incorporates necessary interactions between package and chiplet designs for holistic planning and optimization.
- Our flows can handle for both homogeneous and heterogeneous designs making use of standard ASIC CAD tools.

Future Work

- Extend the holistic/in-context flow with broader technology choices
- Study signal and power integrity with all RCLM elements
- Chiplet-Package co-placement, routing, and optimizations

UNIVERSITY OF ARKANSAS

Thank You

UNIVERSITY OF ARKANSAS

http://csce.uark.edu +1 (479) 575-6043 makabir@uark.edu