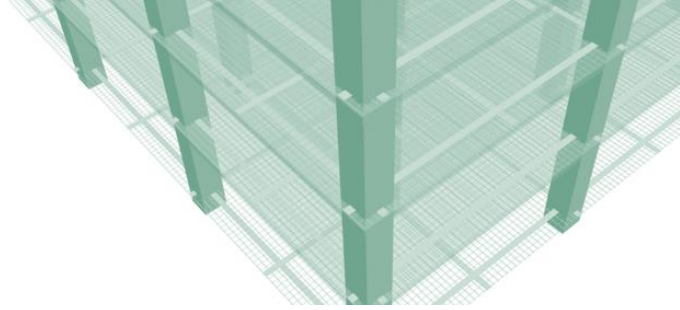
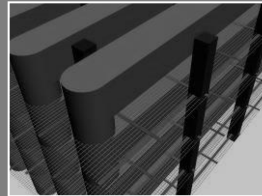
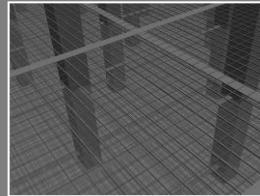
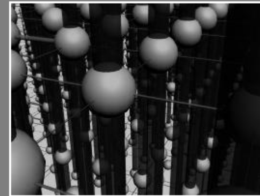
Georgia Institute of Technology

SRC

Mentor Graphics



Full-chip Inter-die Parasitic Extraction in Face-to-Face-Bonded 3D ICs

Yarui Peng[1], Taigon Song[1], Dusan Petranovic[2], and Sung Kyu Lim[1]

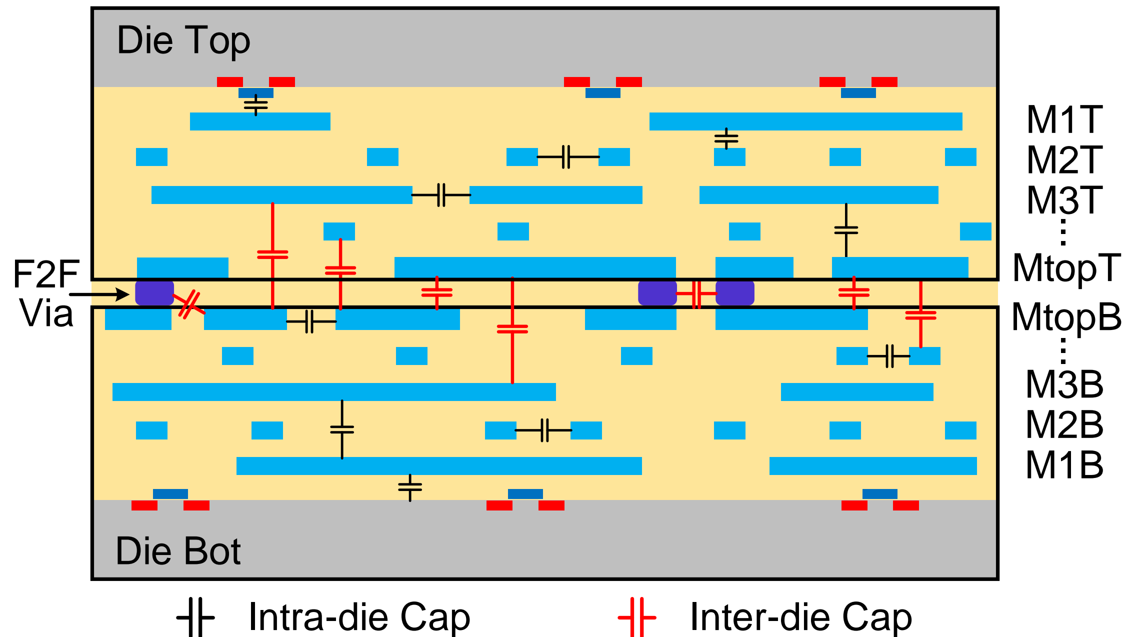
Georgia Institute of Technology, Atlanta, GA, USA

[1]School of ECE, Georgia Institute of Technology, Atlanta, GA, USA

[2]Mentor Graphics, Fremont, CA, USA

Face-to-face Bonding Structure

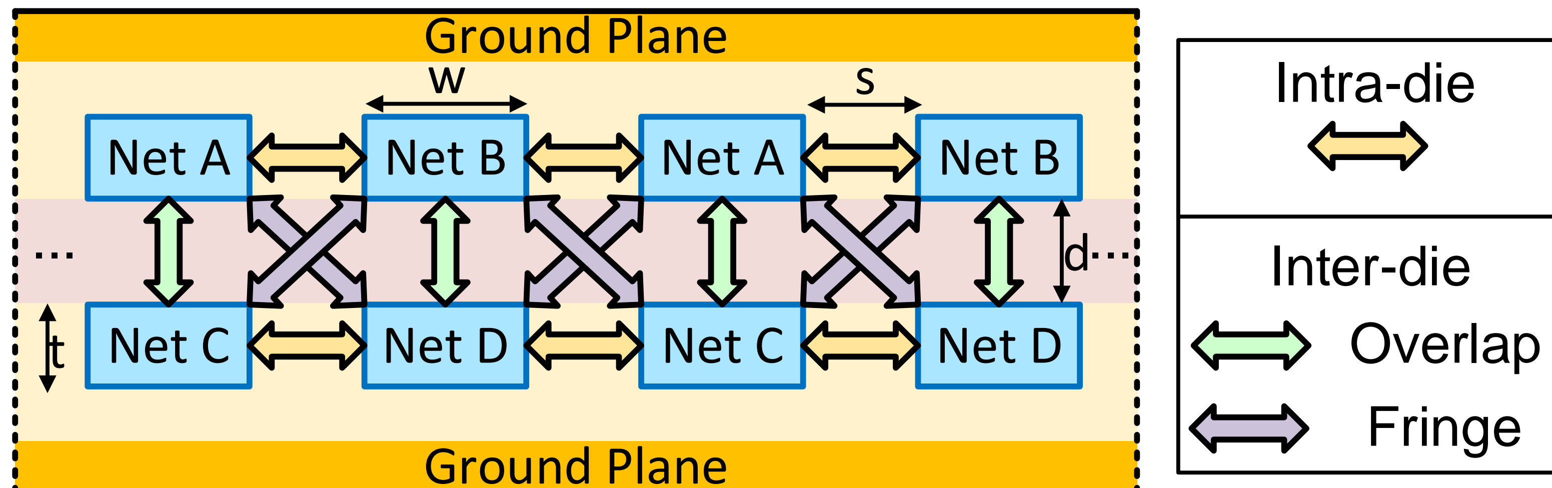
- **Inter-die capacitance becomes important when die-to-die distance is small, especially for face-to-face (F2F) bonded structures with direct copper bonding**



F2F-bonded 3D IC structure with interconnect parasitics

Raphael Simulation Structure

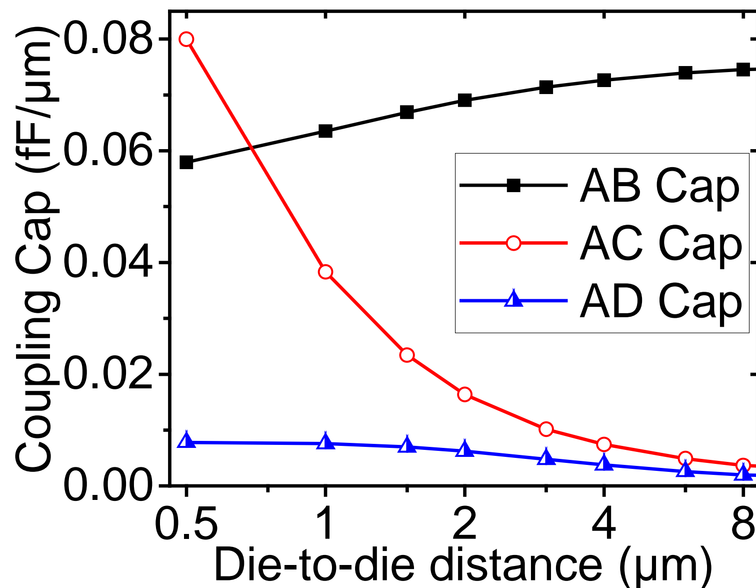
- To analyze the trend in F2F structure, we build a test structure in Raphael with repeated pattern
  - Wire dimensions are based on M4 dimensions in a 45nm technology
  - Intra-die coupling: AB cap and CD cap
  - Inter-die overlap coupling: AC cap and BD cap
  - Inter-die fringe coupling: AD cap and BC cap

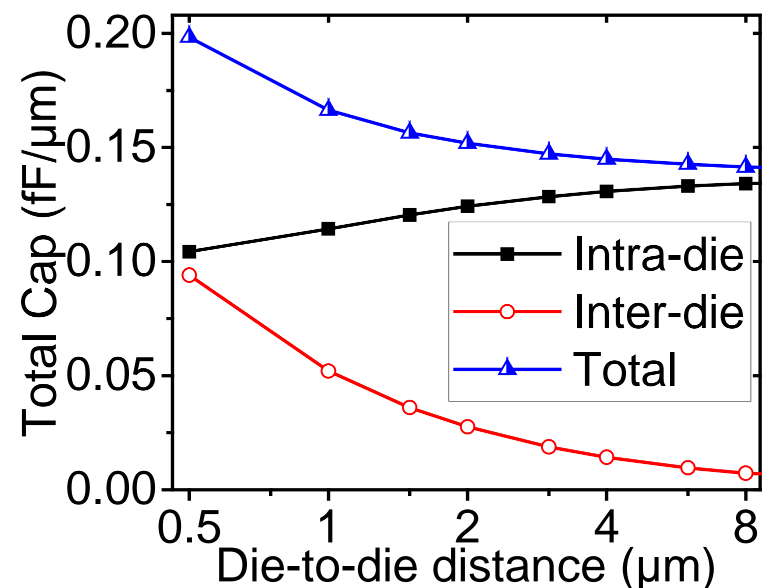


Raphael simulation structure

Die-to-die Distance Impact

- **With a closer die-to-die distance:**
  - **Intra-die cap (AB Cap) decreases due to stronger E-field sharing**
  - **Inter-die cap increases significantly**
  - **Inter-die overlap cap (AC cap) increases much more than inter-die fringe cap (AD cap)**

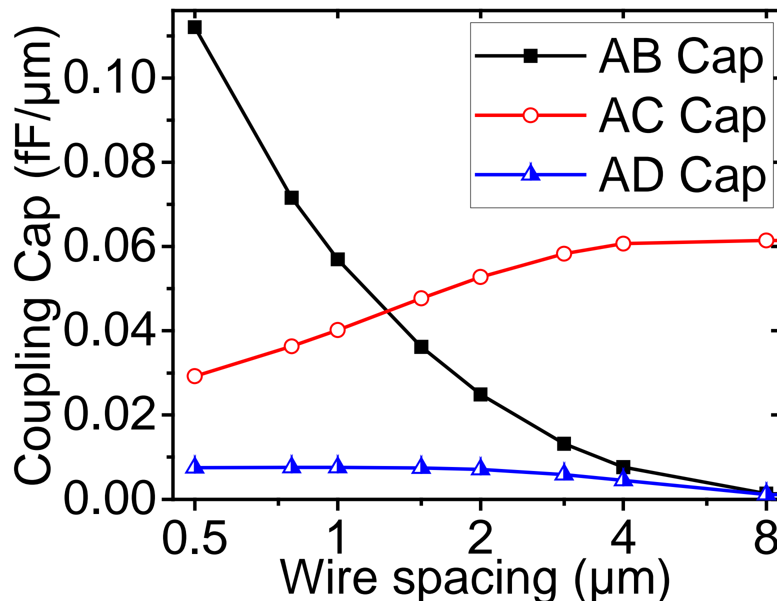


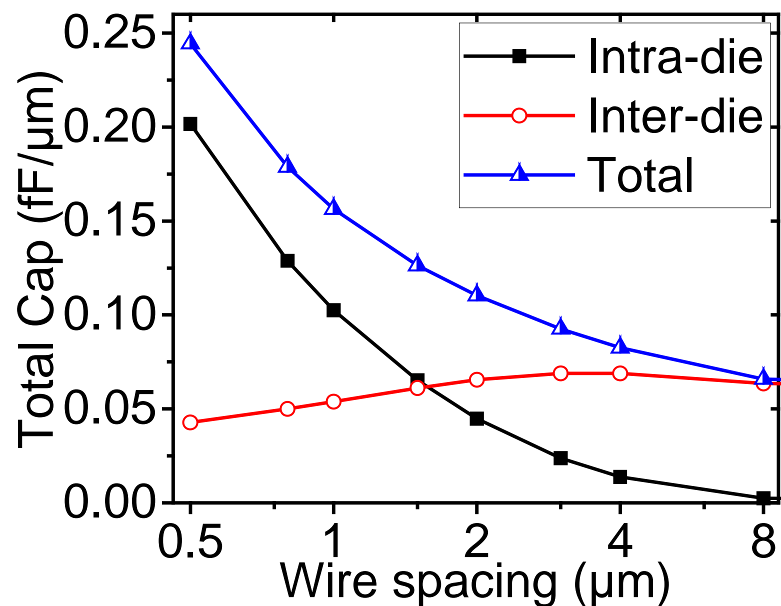


Die-to-die distance impact

Wire Spacing Impact

- **With a larger wire-to-wire distance**
  - **Both intra-die coupling (AB cap) and total cap reduces**
  - **Inter-die first increases with larger overlap cap (AC cap) due to weaker E-field sharing then slightly decrease due to smaller fridge cap (AD cap)**

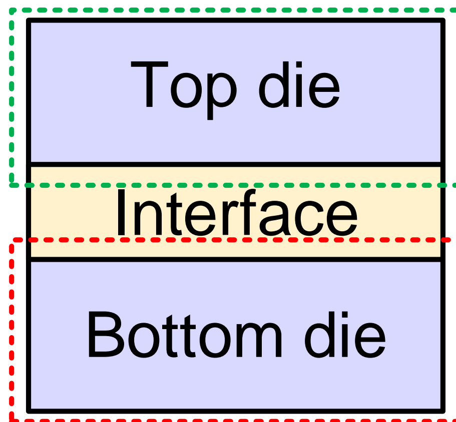




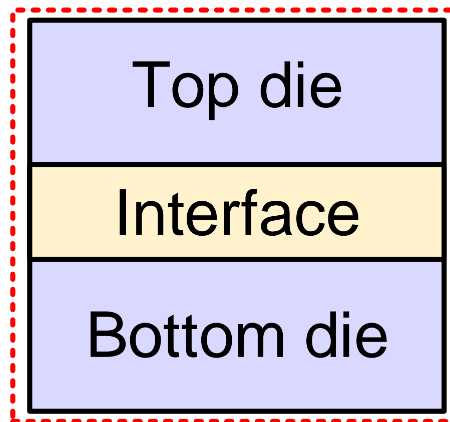
Wire spacing impact

Three Ways of Full-chip F2F Extraction

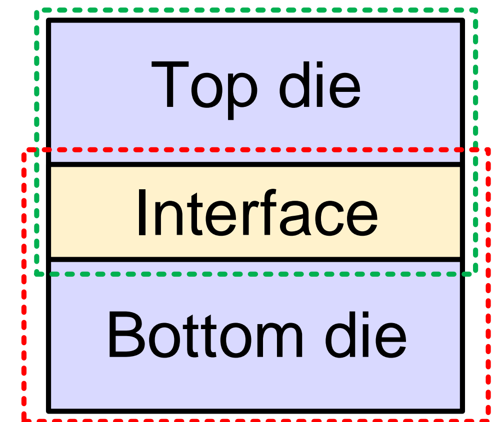
- **Die-by-die extraction**
  - Extract dies separately
- **Holistic extraction**
  - Extract all layers simultaneously
- **In-context extraction**
  - Extract each die separately but aware of a few neighboring die layers



Die-by-die



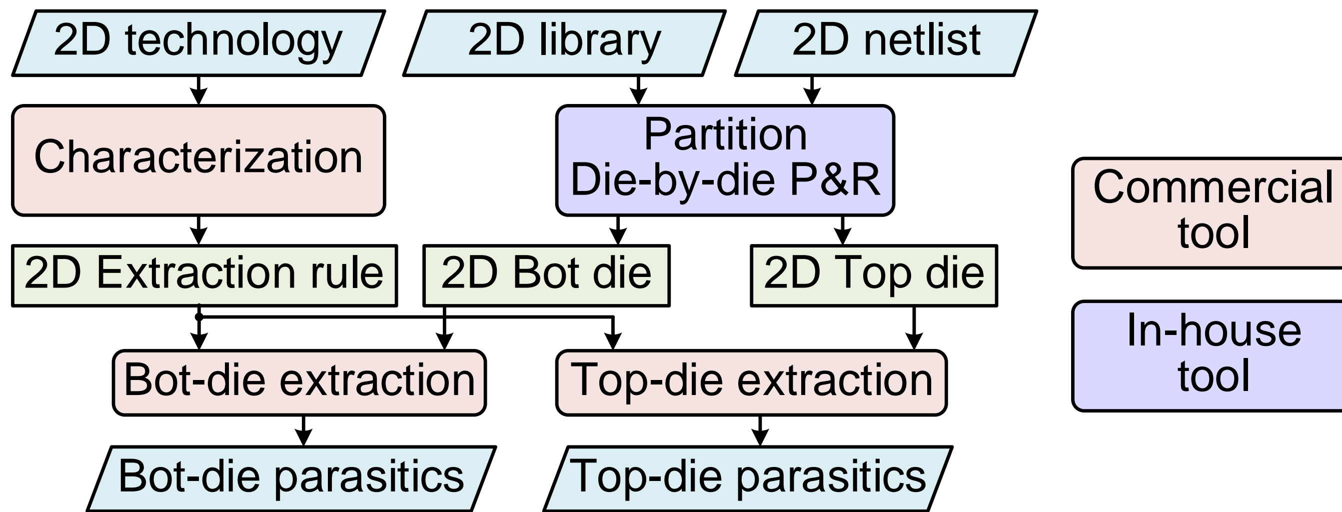
Holistic



In-context

Die-by-die Extraction Flow

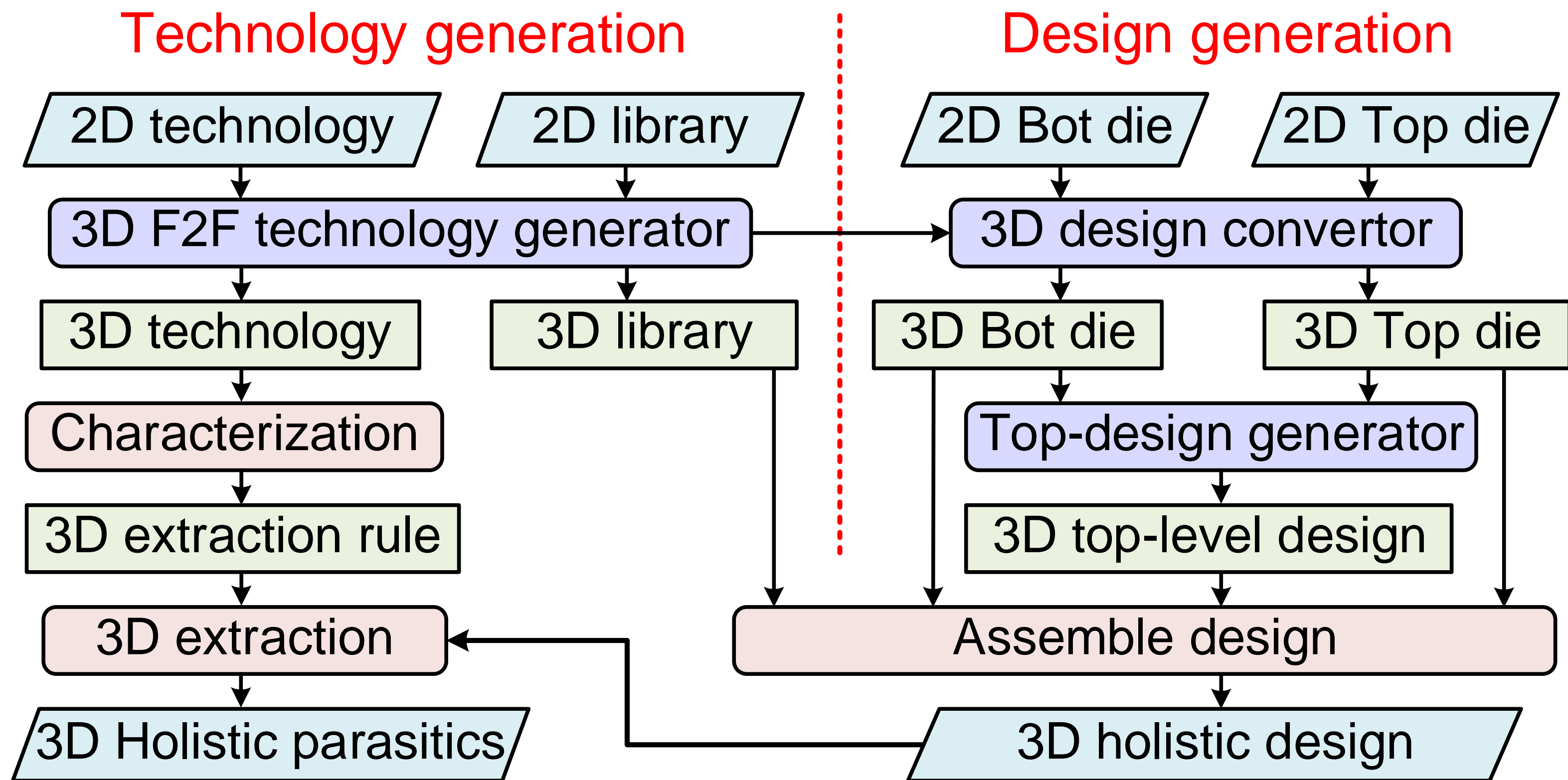
- **Die-by-die extraction is the straight-forward flow currently enabled by many CAD tools**
  - **Assumes each die can be extracted separately**
  - **Ignores all parasitic between dies**
  - **Accurate when dies are separated far or have a ground layer in between**



Die-by-die extraction flow

Holistic Extraction Flow

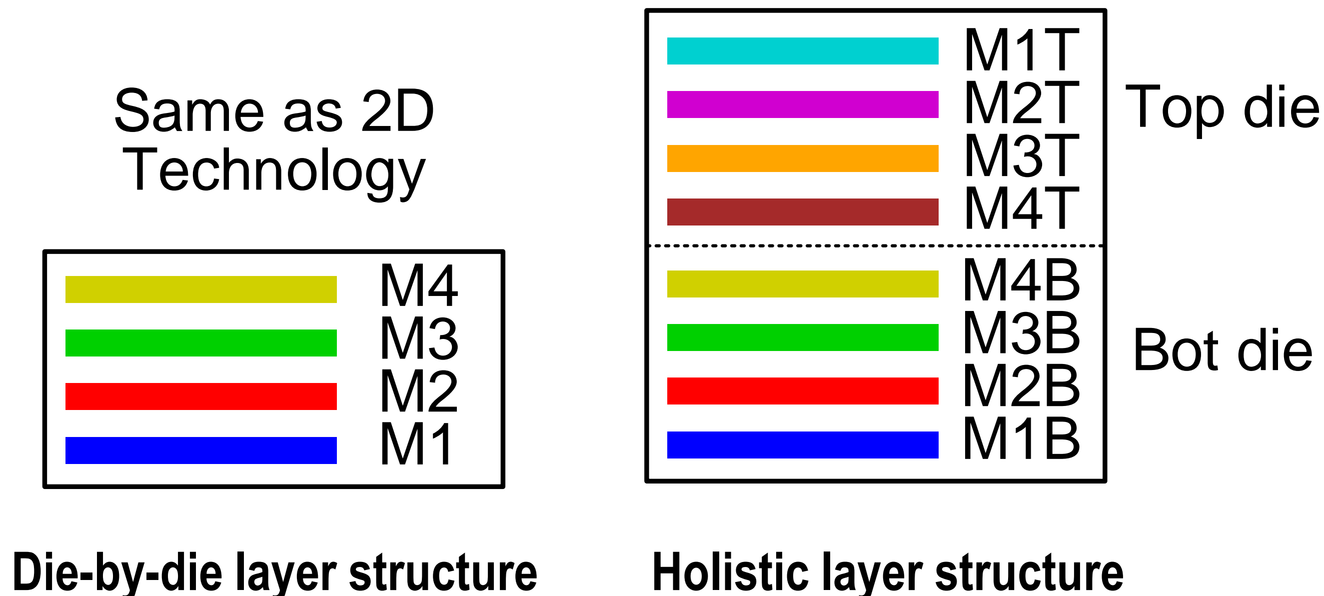
- **Holistic extraction takes all layers into consideration and it introduces more CAD and LVS complexity**



Holistic extraction flow

Die-by-die vs. Holistic Extraction

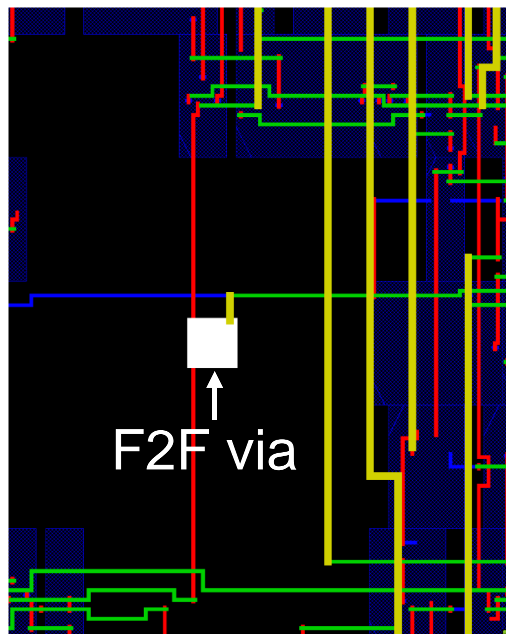
- **Die-by-die uses the same metal stack as 2D technology**
  - **Enables reuse of existing DRC, LVS and PEX rule decks**
- **Holistic extraction needs to rebuild rule decks**
  - **All original and derived layers and device renamed and remapped**
  - **Need technology recalibration**



Die-by-die layer structure **Holistic layer structure**

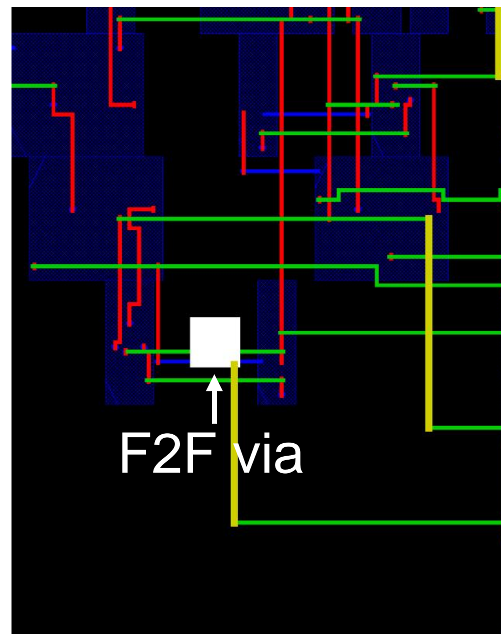
Holistic Design Example

- **By assembling of individual dies, we are able to create a holistic design which contains all metal layers**



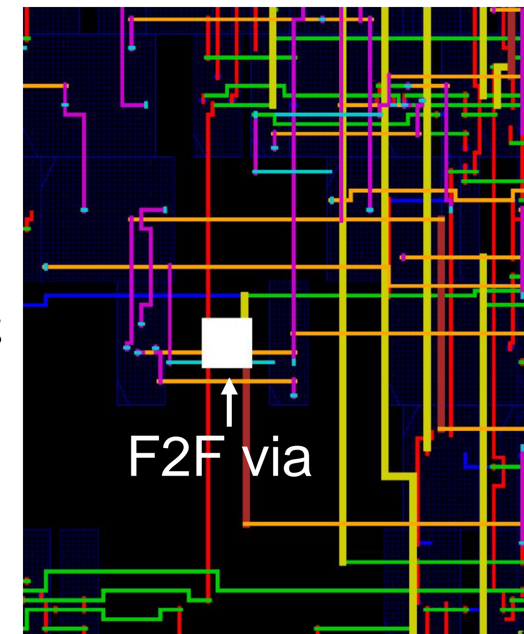
Bottom Die

+



Top Die

=

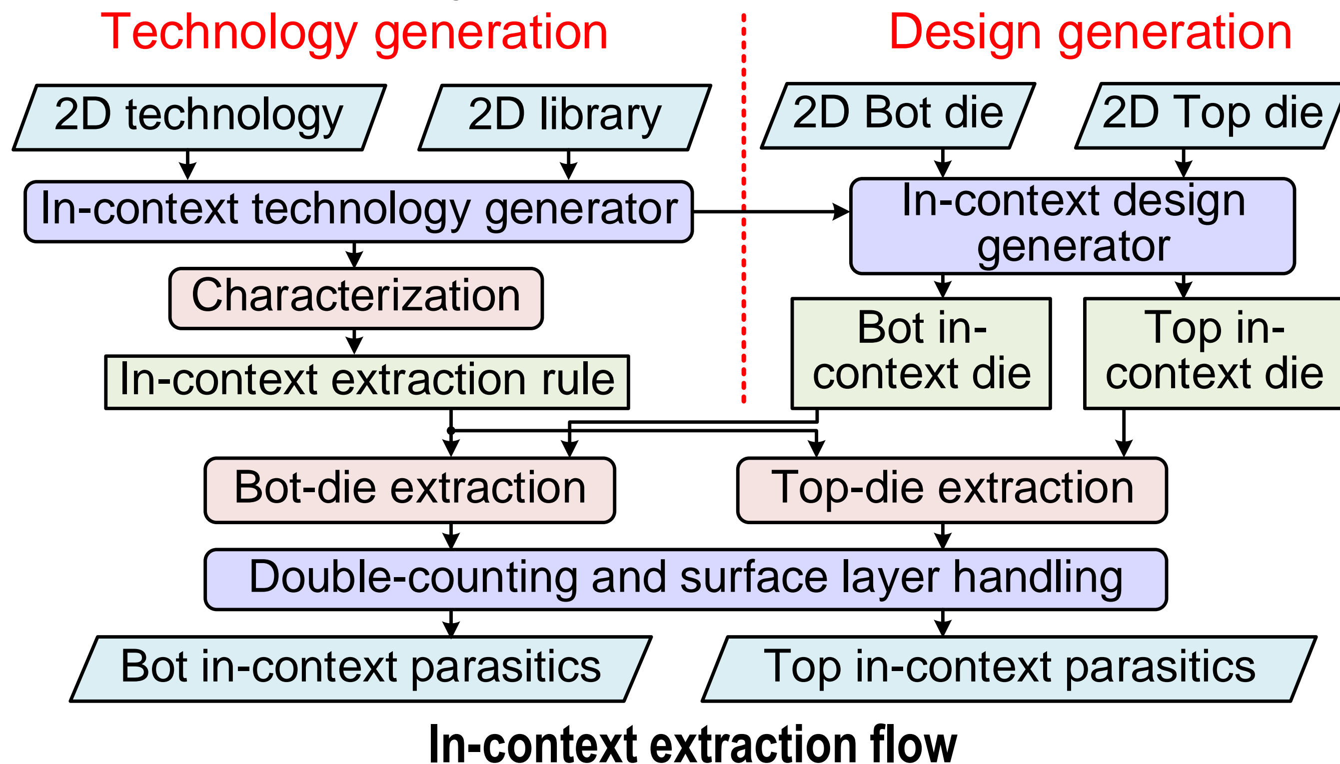


Merged

Design assemble process

In-context Extraction Flow

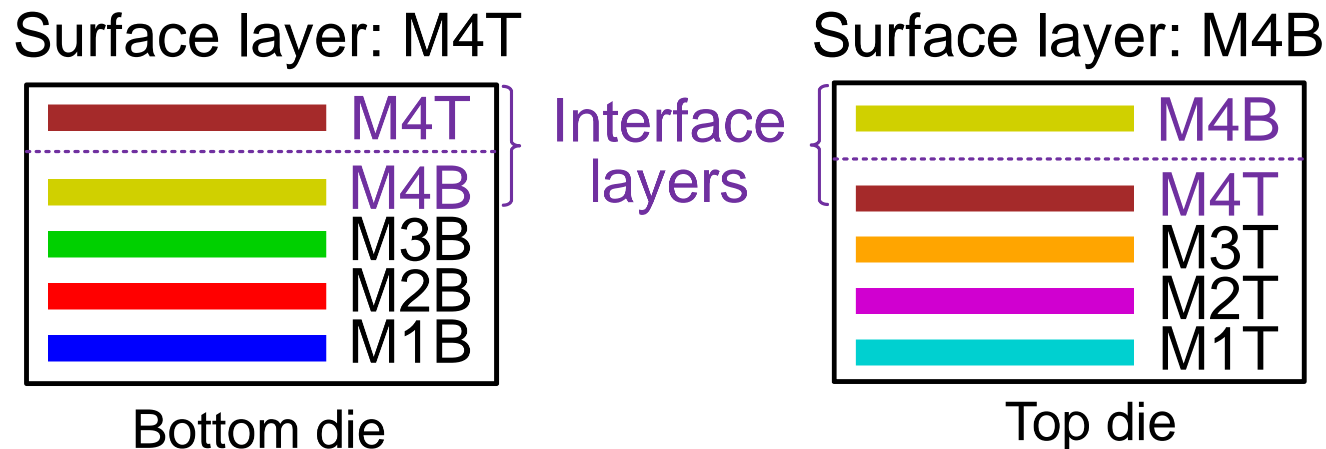
- **In-context extraction takes in a few metal layers from the neighboring die as interface layers**
  - **Keeps most of inter-die coupling and remains accurate**
  - **Reduces CAD complexity and compatible with current tool flow**



In-context extraction flow

In-context Design Example

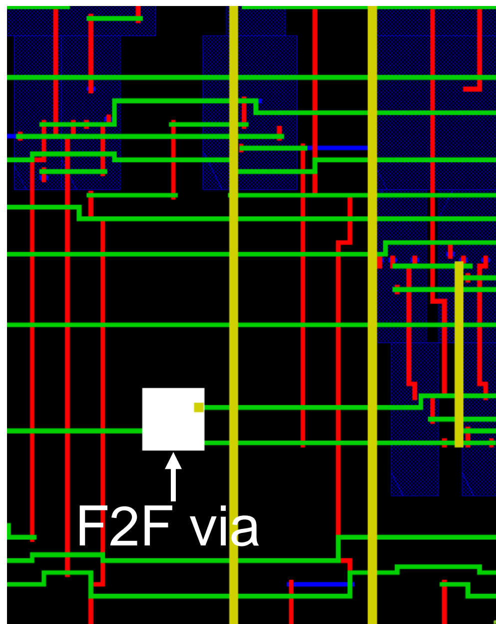
- **In-context technology can be calibrated incrementally**
  - **Base layer calibration results can be derived from existing rule decks**
  - **The surface layer in the in-context extraction is defined as the layer furthest from the substrate**



In-context layer structure
(with one interface layer from each die)

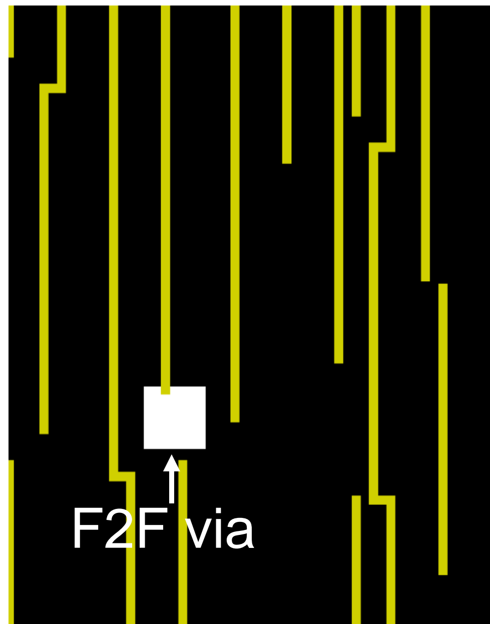
In-context Design Example

- **In-context design only needs additional routing information from the neighbor die**
  - **Enables much simpler rule deck generation**



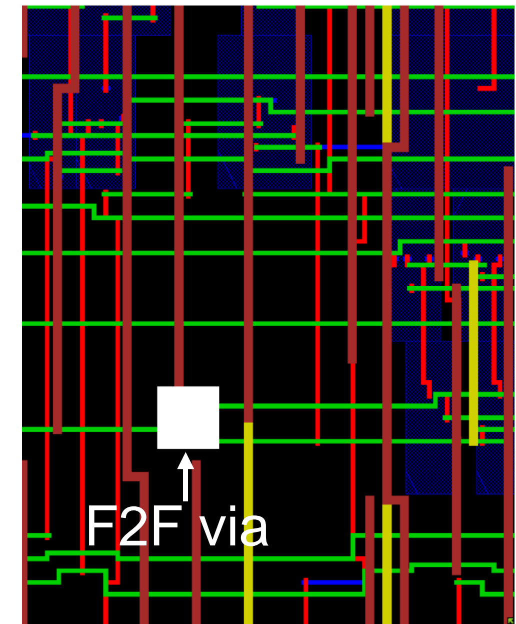
Bottom Die

+



Top Interface layer

=

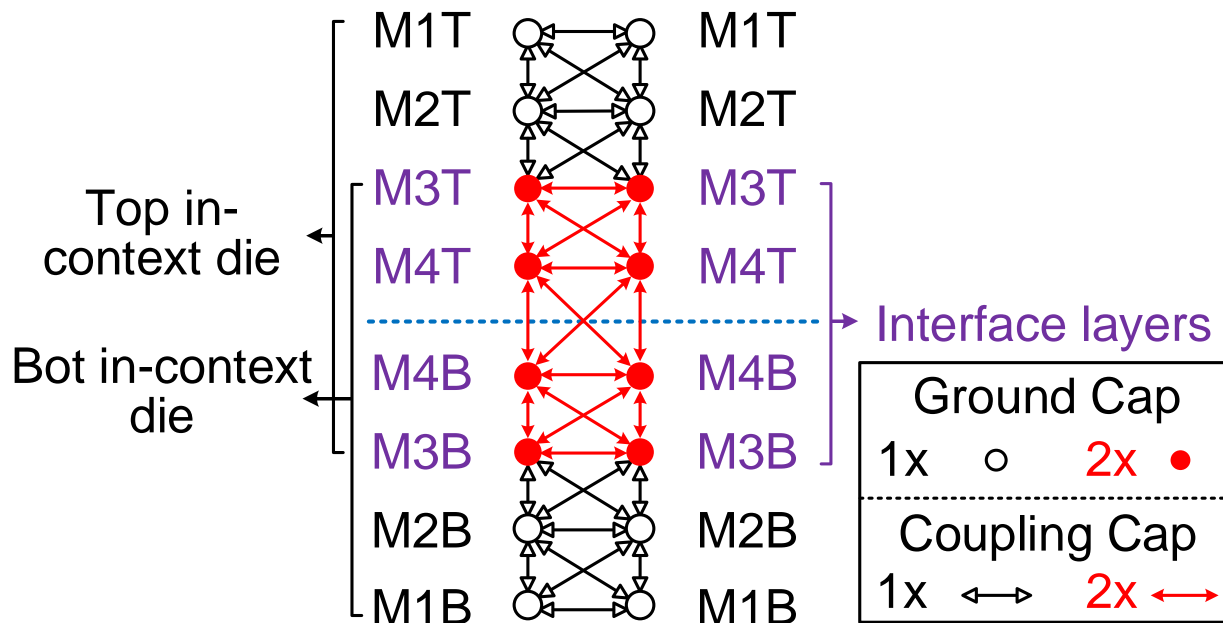


Bottom In-context die

In-context design generation

Double Counting Correction

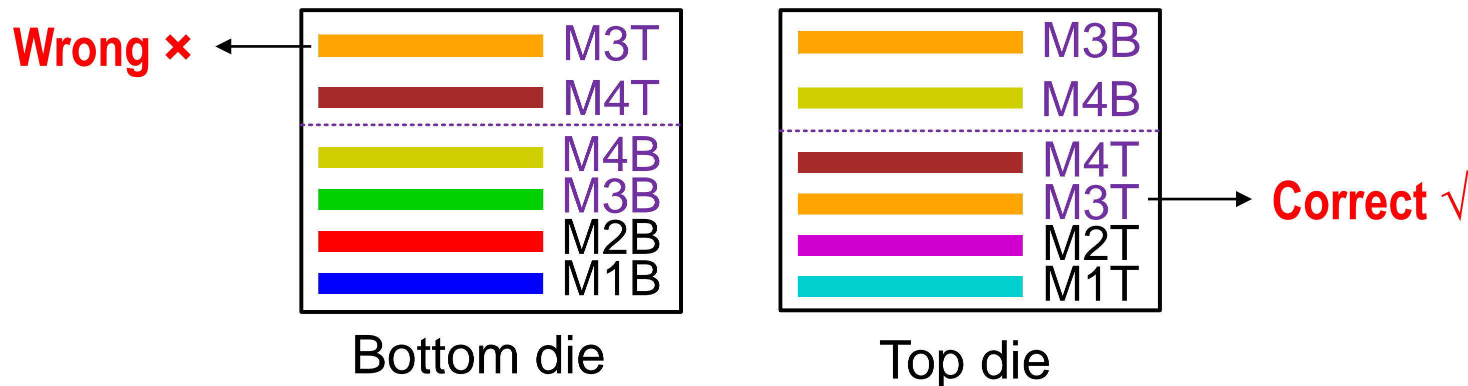
- **With in-context extraction, capacitance on interface layers are double-counted**
  - **A simple solution is to halve all caps from interface layers in SPEF files**



Double-counted capacitance with in-context extraction

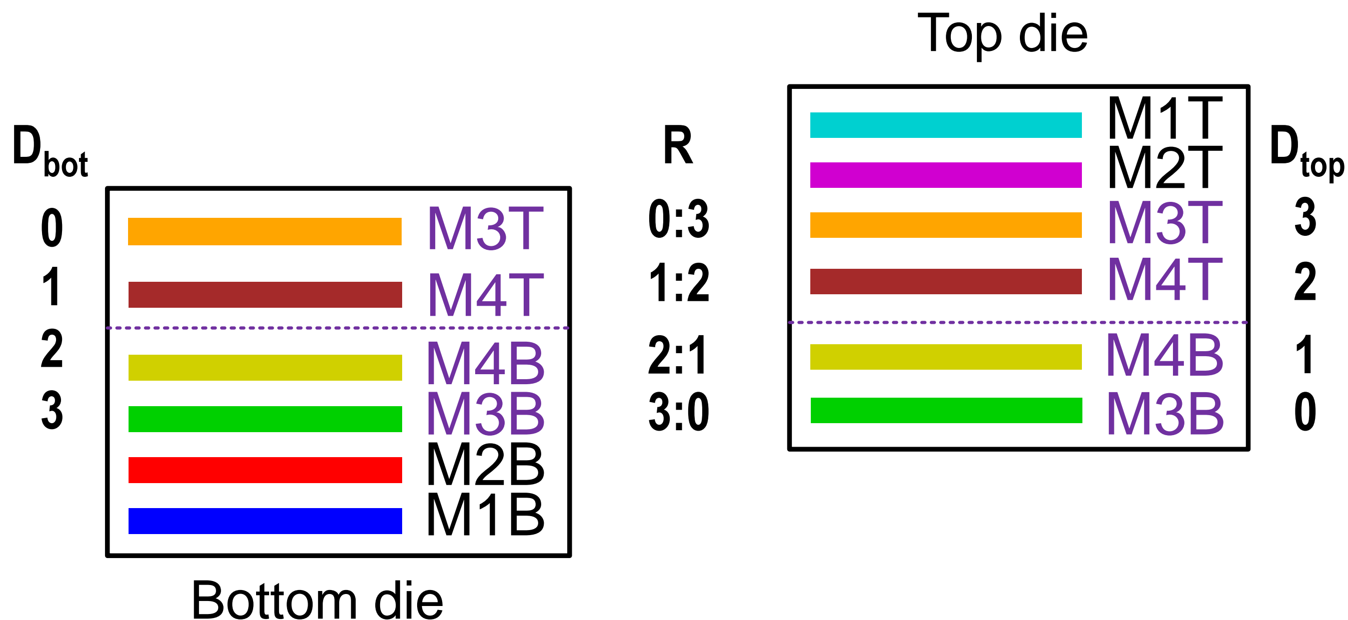
Surface Layer Correction

- **Surface layer only sees one neighboring layer**
  - **Introduce large error with less E-field sharing**
- **Note each layer is not the surface layer in both in-context dies**
  - **E.g., M3T is the surface layer in bottom die but not in top die**
- **Surface layer correction based on weighted average**
  - **Use a weighted average for caps on interface layers**
  - **Larger weight for layers farther from the surface**



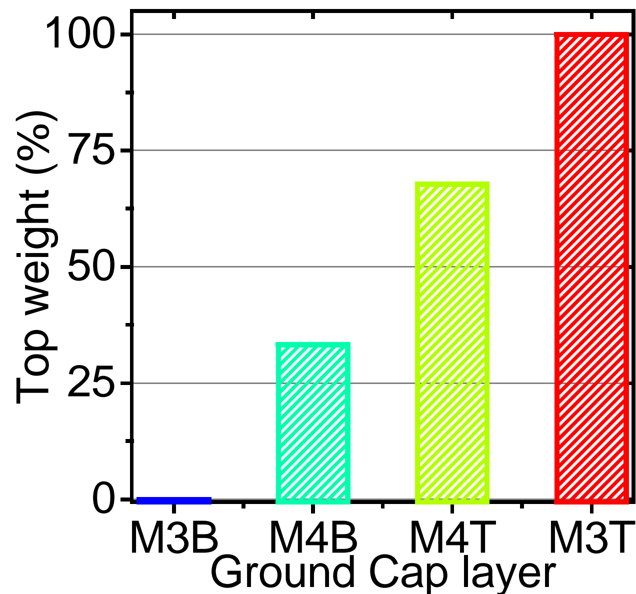
Surface Layer Correction

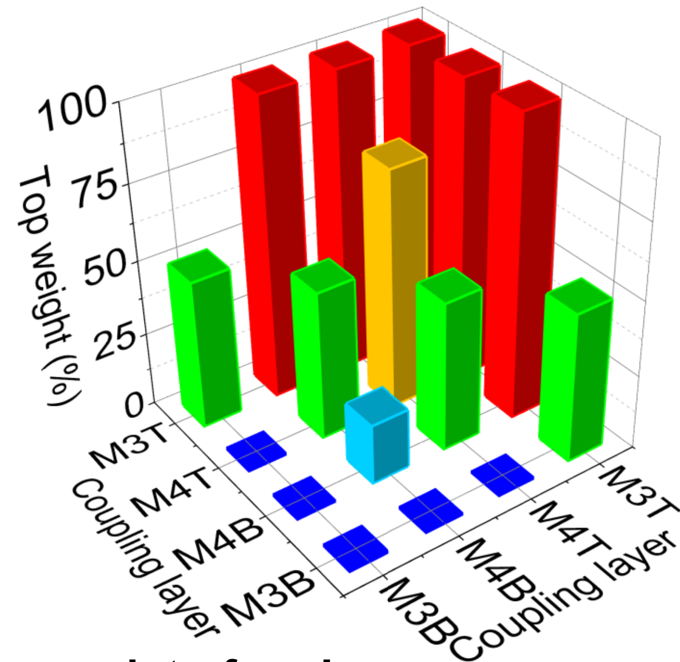
- **For each layer, we define:**
  - **D: distance to the surface**
  - **R: ratio between D values in the bottom and top in-context die**
- **Example (two interface layers per die)**



Surface Layer Correction

- **The surface correction weight of a capacitor is the product of R ratios of both layers it connects to, normalized to 100%**
  - The R ratio of ground layer is defined as 1:1
  - A ground cap on M4T: $R_{M4T} \times R_{gnd}$ = 1:2 × 1:1 = 33% (bot) : 67% (top)
  - A cap between M4T and M4B: $R_{M4T} \times R_{M4T}$ = 1:2 × 2:1 = 50% (bot) : 50% (top)



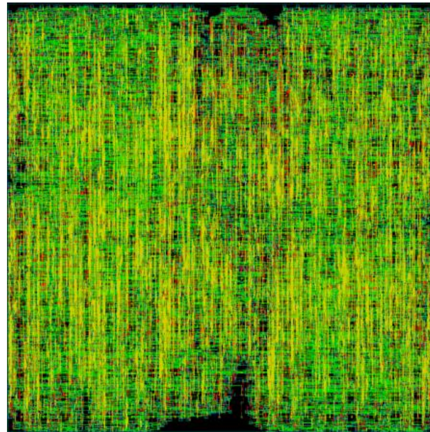


Weighted average for caps on interface layers

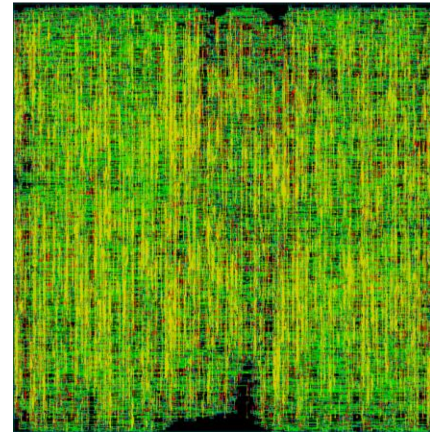
Sample FFT Design in F2F

- **A 64-point FFT with 38K gates and 330 F2F vias is implemented**

Die-by-die

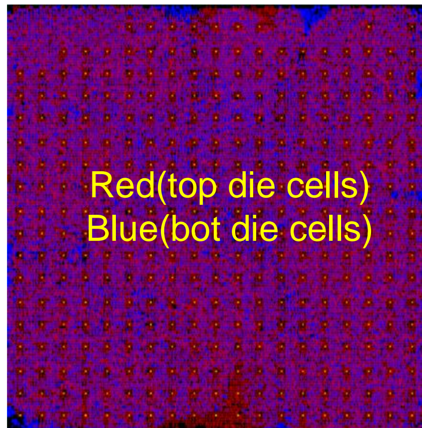


Bottom Die

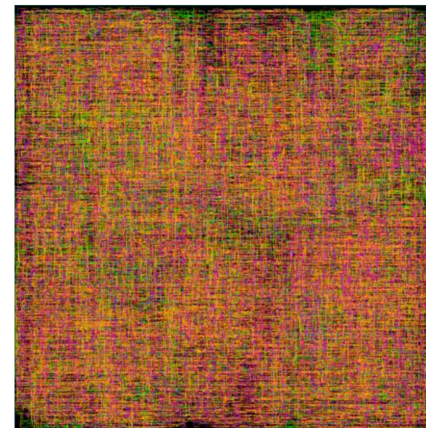


Top Die

Holistic



Placement



Routing

Die-by-die vs. Holistic Extraction

- **With a 1um F2F via height, die-by-die extraction underestimates coupling capacitance significantly**
  - **Especially for layers close to the other die**

Layer	M1B	M2B	M3B	M4B	M4T	M3T	M2T	M1T	Total
Holi	26.2	949	1808	3703	3089	1755	1013	38.2	12381
D-D	20.1	856	1620	1955	1413	1399	747	21.2	8032
Err	-6.06	-93.4	-187	-1747	-1676	-356	-266	-17.0	-4349
Err%	-23%	-9.8%	-10%	-47%	-54%	-20%	-26%	-45%	-35%

Total coupling capacitance of each layer

Breakdown of Holistic Extraction

- **Inter-die coupling occupies a large portion of total coupling cap**
  - **Especially when dies are close and few metal layers are used**

Layer	M1B	M2B	M3B	M4B	M4T	M3T	M2T	M1T
M1B	5.76	3.03	17.1	0.13	0.03	0.14	0.00	0.00
M2B	3.03	381	147	396	18.6	0.69	2.58	0.01
M3B	17.1	147	1261	231	9.9	140	0.72	0.28
M4B	0.13	396	231	1826	1184	18.6	46.7	0.12
M4T	0.03	18.6	9.9	1184	1311	196	369	0.28
M3T	0.14	0.69	140	18.6	196	1226	148	25.3
M2T	0.00	2.58	0.72	46.7	369	148	442	4.63
M1T	0.00	0.01	0.28	0.12	0.28	25.3	4.63	7.54

Breakdown of holistic extraction

In-context vs. Holistic Extraction

- **Our in-context extraction with double counting and surface layer corrections matches very well with holistic extraction**
  - Using two interface layers from each die

Layer	Ground capacitance								
	M1B	M2B	M3B	M4B	M4T	M3T	M2T	M1T	Total
Holi	1136	6588	9240	3878	2664	8320	6306	1117	39247
In-C	1137	6583	9249	4159	2639	8183	5986	949	38886
Err	1.10	-4.20	9.00	281	-24.9	-136	-319	-168	-361
Layer	Coupling capacitance								
	M1B	M2B	M3B	M4B	M4T	M3T	M2T	M1T	Total
Holi	26.2	949	1808	3703	3089	1755	1013	38.2	12381
In-C	26.3	950	1803	3679	3058	1734	1001	38.0	12287
Err	0.15	0.81	-5.15	-24	-31.0	-21.3	-12.3	-0.22	-93.3

Interface Layer Impact

- **More interface layers helps improve accuracy**
  - **With two interface layers per die gives a good tradeoff**

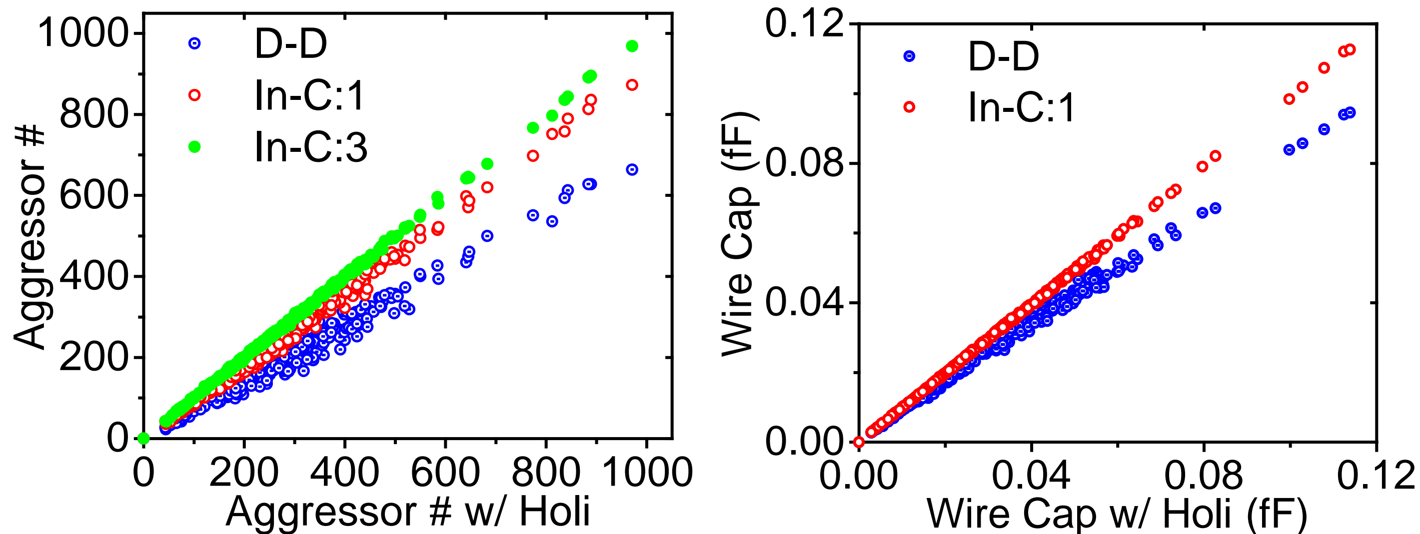
Layer	M1B	M2B	M3B	M4B	M4T	M3T	M2T	M1T	Total	Err%
Holistic	26.2	949	1808	3703	3089	1755	1013	38.2	12381	-
In-C:1	26.1	953	1701	3708	2994	1604	994	37.8	12018	-2.93%
In-C:2	26.3	950	1803	3679	3058	1734	1001	38.0	12287	-0.76%
In-C:3	26.2	949	1794	3671	3057	1745	1012	38.2	12292	-0.72%

- **Our weighted methods improves in-context extraction accuracy**

Layer	M3B	M4B	M4T	M3T	Total
Holi	1808	3703	3089	1755	10354
Original	3069	6779	5781	3522	19151
Halved	1618	3611	3082	1849	10159
Weighted	1803	3679	3058	1734	10272

In-context Extraction Accuracy

- **In-context extraction captures inter-die aggressors, provides better accuracy in full-chip analysis**
  - **Especially for 3D nets which communicates across dies**



In-context extraction of 3D nets

Full-chip Analysis Results

- **Full-chip analysis also shows non-negligible impact from inter-die capacitance, especially on noise results and 3D nets**
  - **Die-by-die extraction underestimates delay, power and noise**
  - **In-context extraction gives much more accurate results**

Primetime measurement	Holi	D-D	Err%	In-C	Err%
Longest path delay (ns)	3.90	3.66	-6.2%	3.81	-2.3%
3D nets switching power (mW)	1.05	1.01	-3.5%	1.04	-0.5%
Total switching power (mW)	12.1	11.9	-1.7%	12.0	-0.8%
Total coupling cap on 3D nets (fF)	4.37	2.96	-32%	4.19	-4.1%
Total wire cap on 3D nets (fF)	10.8	9.35	-13%	10.6	-1.8%
Avg aggressor # on 3D nets	285	200	-30%	277	-2.8%
Max noise on 3D nets (mV)	41.3	30.40	-26%	38.8	-6.1%

Conclusions

- We studied impacts of E-field sharing in F2F structure
- We showed inter-die coupling cannot be ignored in F2F-bonded 3D ICs, especially with few metal layers and close die-to-die distance
- We implemented and compared three extraction methods with full-chip analysis results
  - Die-by-die extraction underestimates total coupling capacitance
  - Holistic extraction is able to capture all inter-die coupling at the cost of high complexity
  - Our first-of-its-kind in-context extraction is highly accurate, and captures most E-field interactions across dies