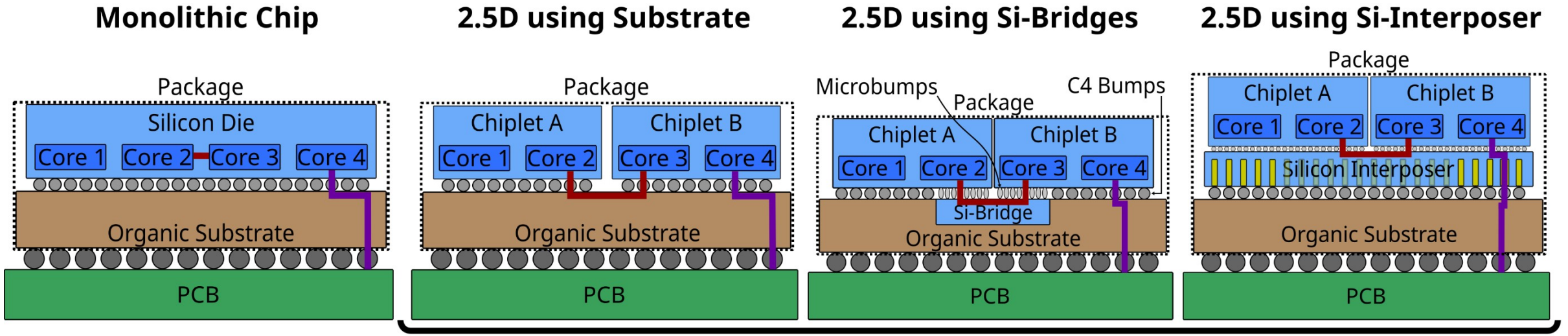


PATRICK IFF, BENIGNA BRUGGMANN, MACIEJ BESTA, LUCA BENINI, TORSTEN HOEFLER

PlaceIT: Placement-Based Inter-Chiplet Interconnect Topologies



Background on 2.5D Integration



Advantages

- Heterogeneity
- Chiplet Reuse
- Improved Yield
- Per-Chiplet Power-Binning

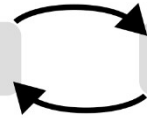
Challenges

- Increased Complexity
- PHY Overheads
- Increased Inter-Chiplet Latency
- Reduced Inter-Chiplet Throughput

Motivation

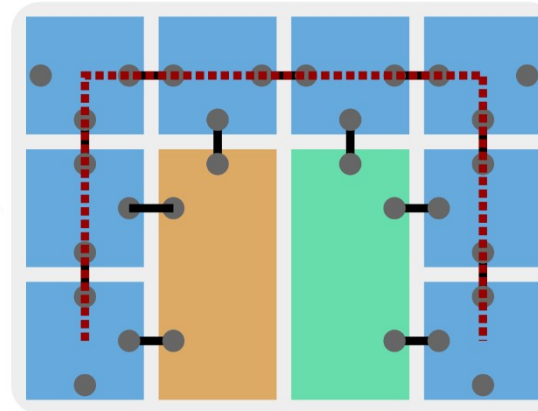
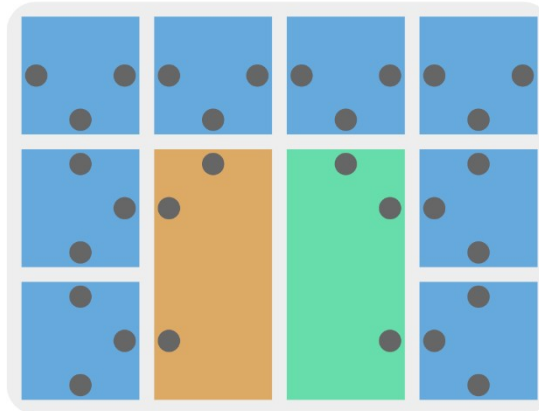
Interconnect Latency and Throughput depend on:

Chiplet Placement

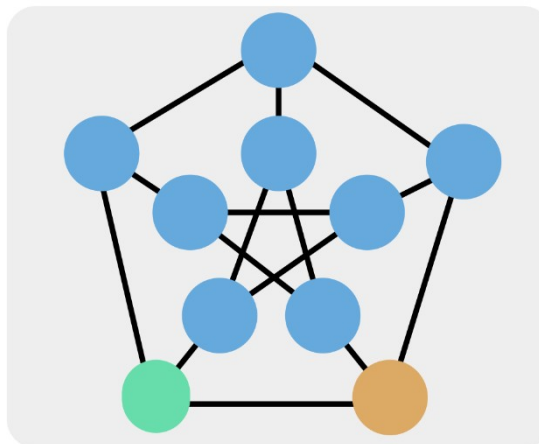


Interconnect Topology

Co-Optimization



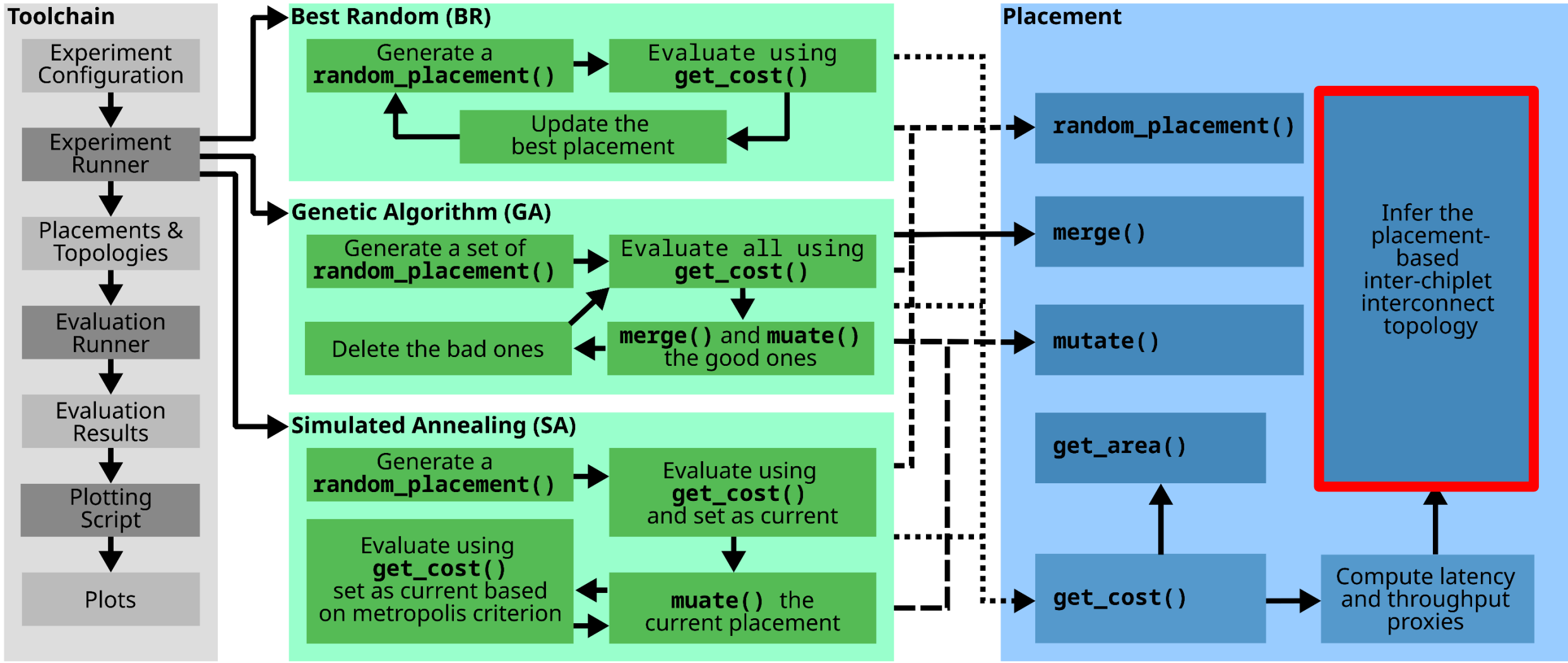
Placement first
Topology second



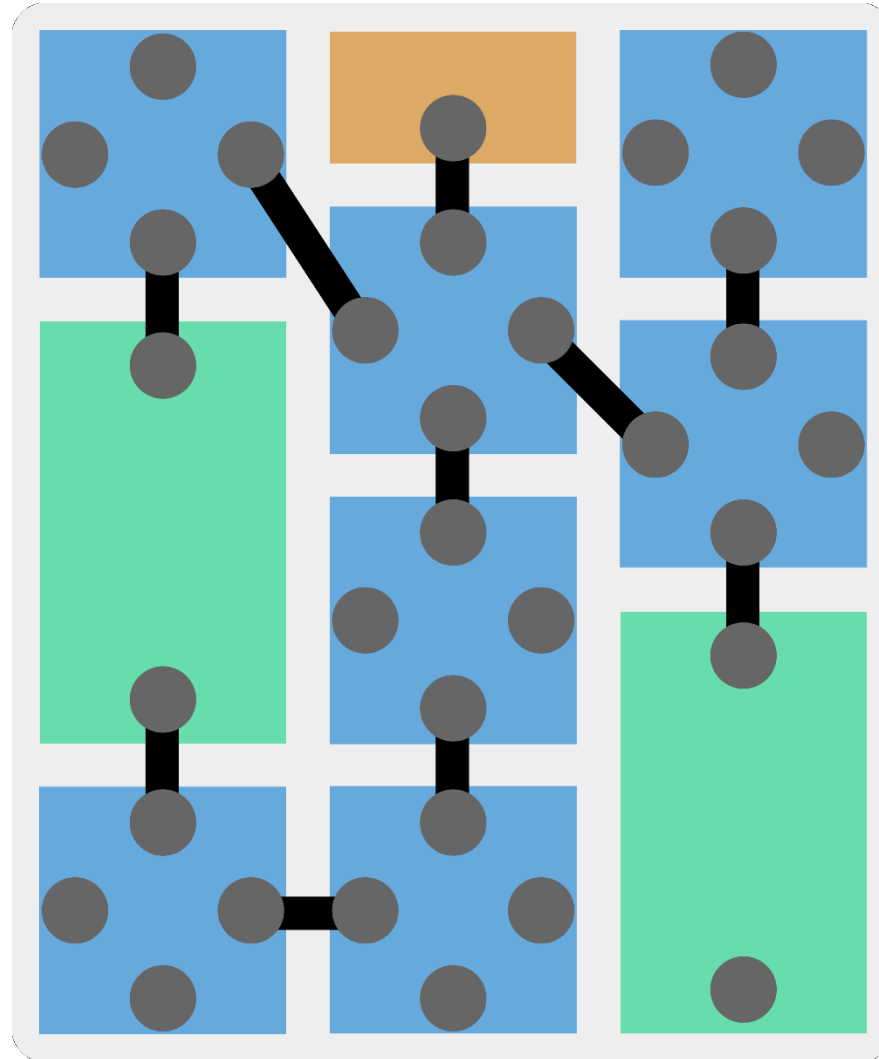
Topology first
Placement second



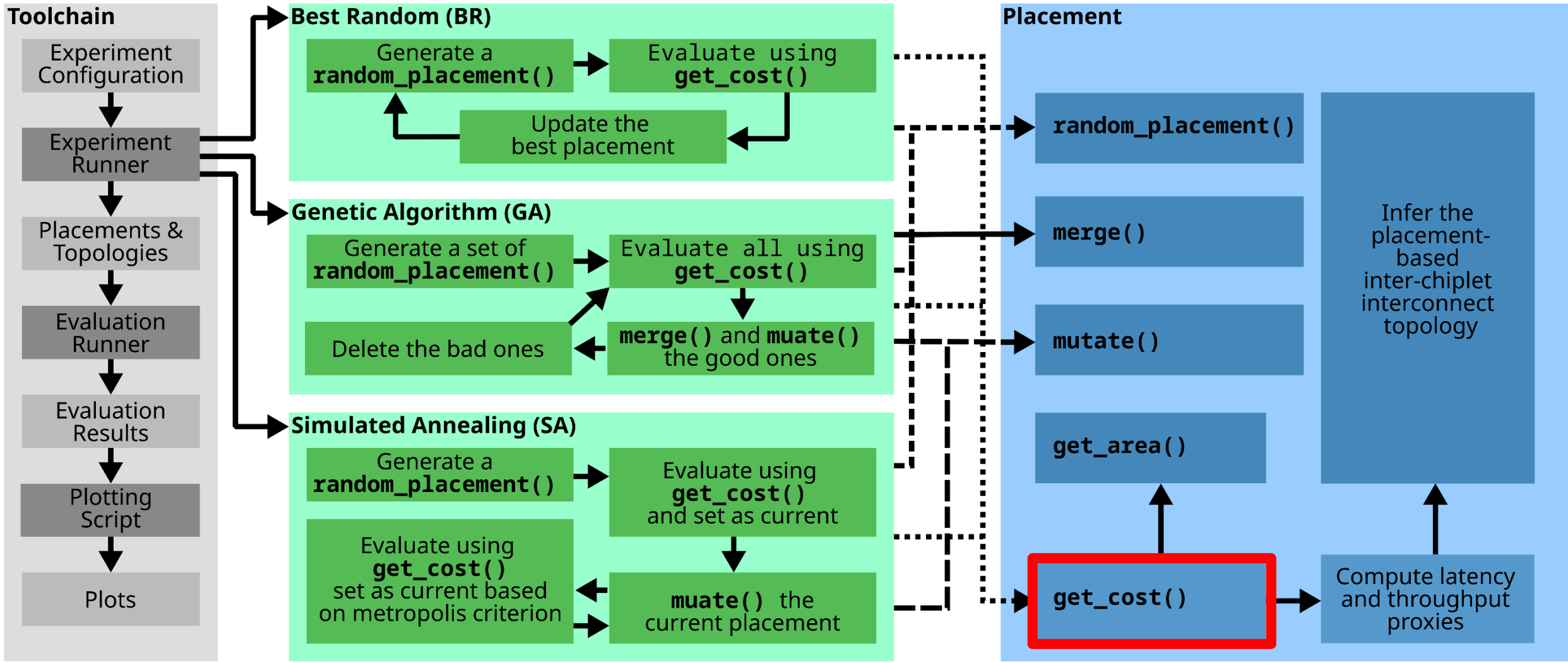
PlaceIT: Architecture



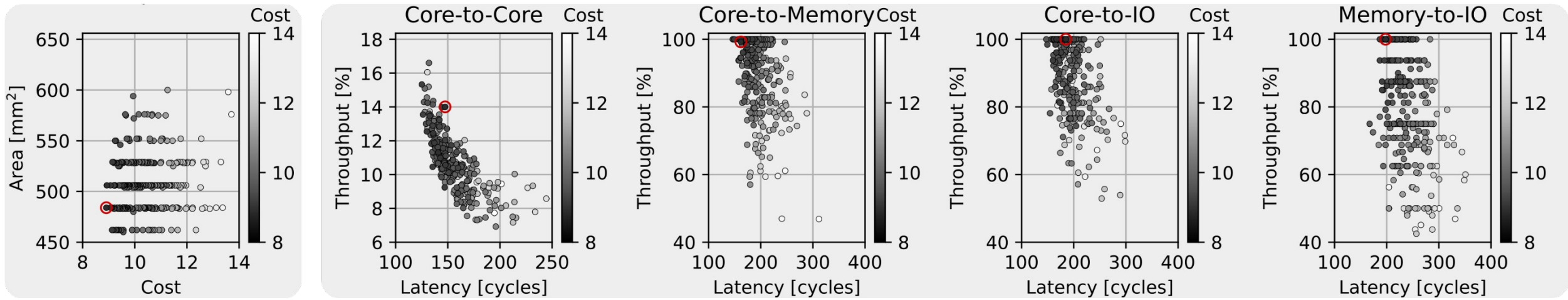
PlaceIT: Inferring Placement-Based Interconnect Topologies



PlaceIT: Architecture

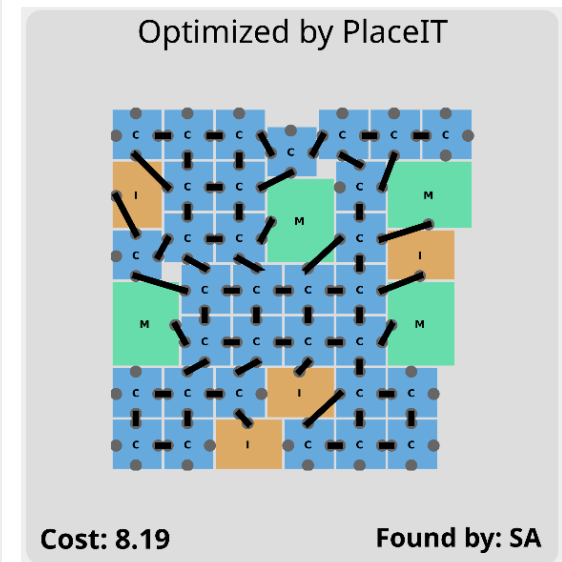
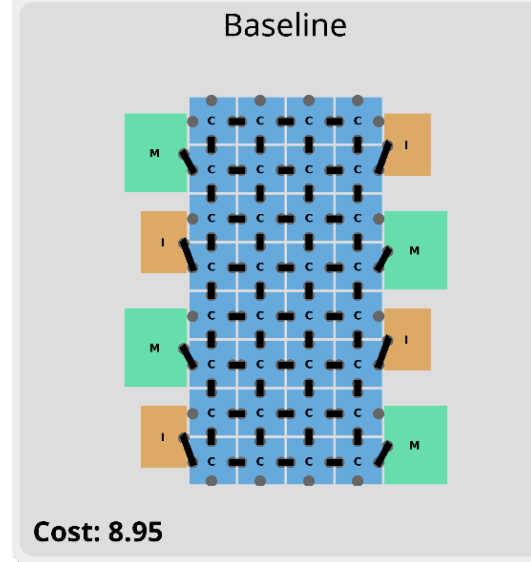
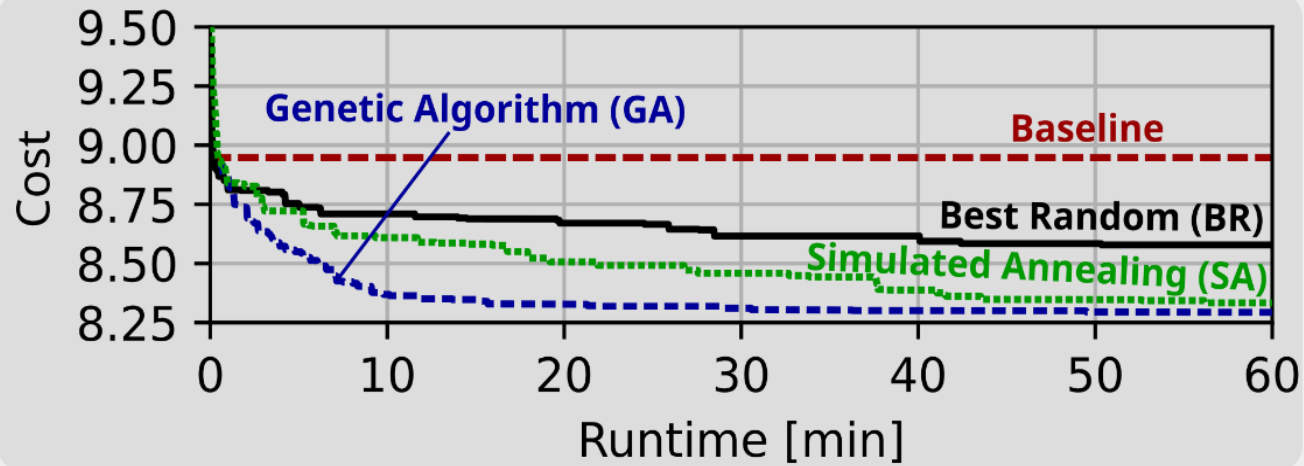


PlaceIT: Cost Function

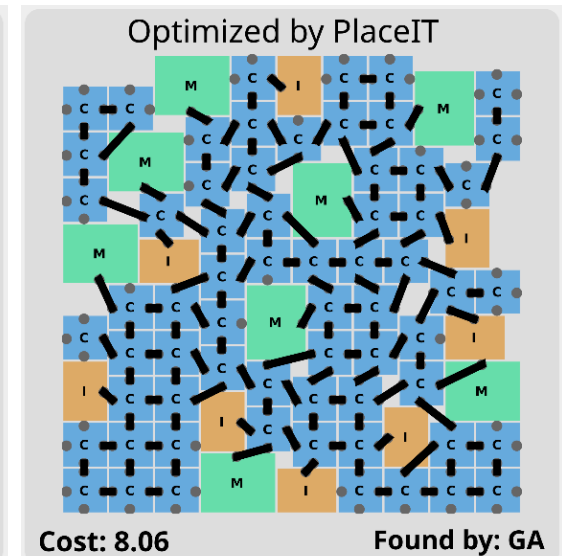
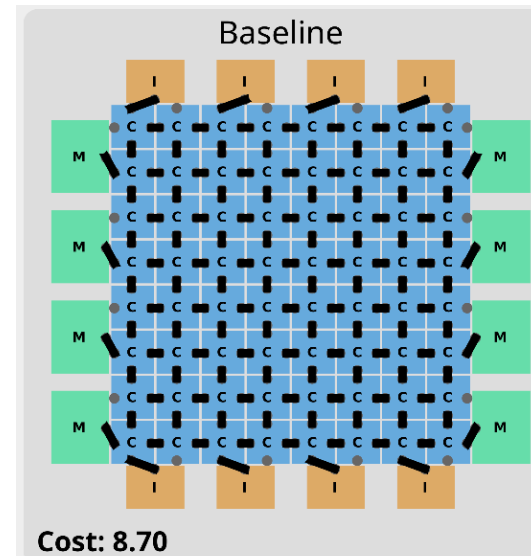
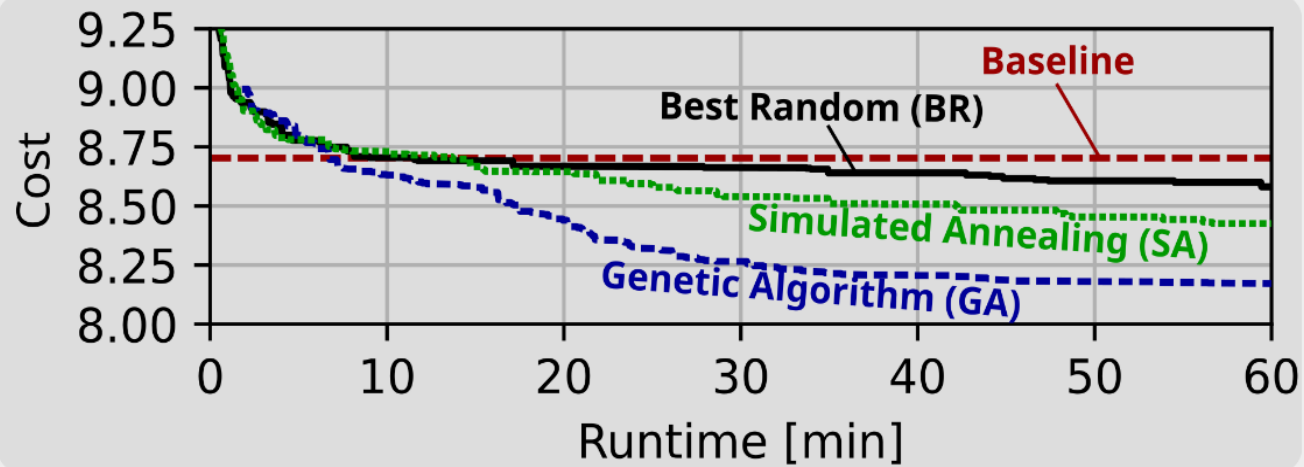


Optimizing Chips with Heterogeneously Shaped Chiplets

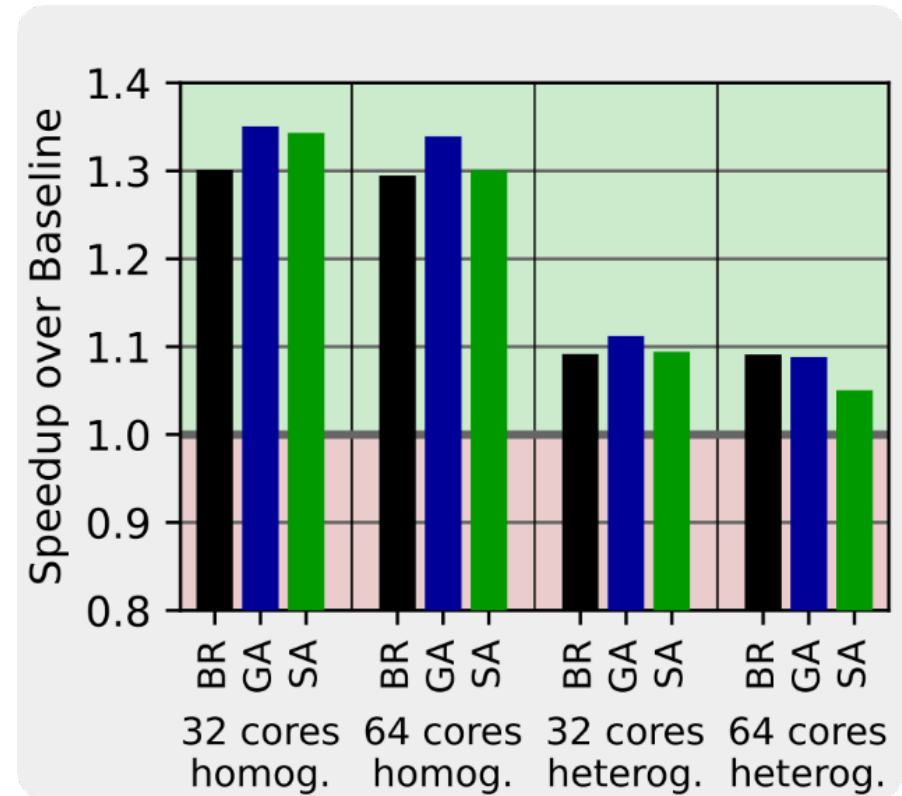
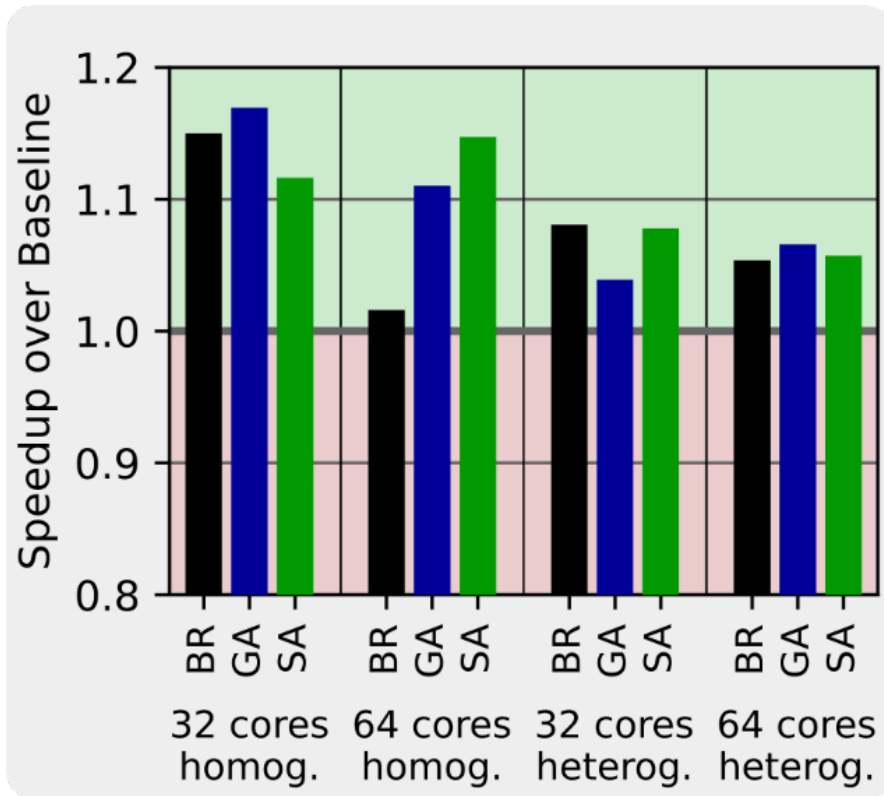
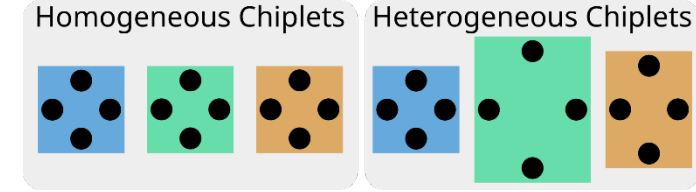
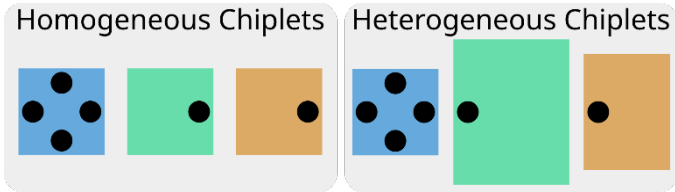
32 cores



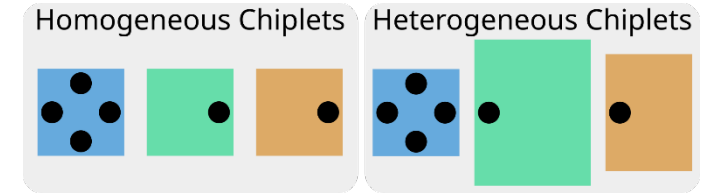
64 cores



Evaluation using Full Traffic Trace



Evaluation using Partial Traffic Trace Regions



Homogeneous

32 cores

	Region 1			Region 2			Region 3			Region 4			Region 5		
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA
blackscholes	92%	86%	88%	77%	79%	83%	93%	85%	82%	95%	96%	99%	109%	101%	98%
bodytrack	90%	86%	89%	77%	79%	83%	76%	75%	69%	88%	81%	81%	101%	90%	90%
canneal	90%	86%	89%	77%	79%	83%	98%	90%	89%	95%	85%	86%	104%	87%	87%
dedup	92%	86%	88%	77%	79%	83%	157%	119%	122%	87%	86%	92%	75%	81%	87%
ferret	92%	86%	88%	77%	79%	83%	102%	94%	94%	95%	92%	88%	87%	74%	86%
fluidanimate	92%	86%	88%	77%	79%	83%	88%	83%	85%	85%	78%	82%	93%	82%	83%
swaptions	90%	86%	89%	77%	79%	83%	96%	91%	93%	84%	85%	81%	110%	101%	102%
vips	92%	86%	88%	77%	79%	83%	100%	92%	91%	80%	75%	78%	78%	68%	74%
x264	90%	86%	89%	77%	79%	83%	81%	71%	71%	87%	84%	91%	108%	98%	96%

64 cores

	Region 1			Region 2			Region 3			Region 4			Region 5		
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA
blackscholes	96%	93%	95%	82%	103%	102%	98%	88%	91%	106%	95%	91%	95%	93%	98%
bodytrack	94%	94%	97%	82%	103%	102%	167%	110%	103%	96%	87%	91%	91%	83%	86%
canneal	94%	94%	97%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
dedup	96%	93%	95%	82%	103%	102%	97%	95%	107%	97%	93%	87%	90%	90%	104%
ferret	96%	93%	95%	82%	103%	102%	90%	87%	94%	97%	83%	88%	93%	82%	88%
fluidanimate	96%	93%	95%	82%	103%	102%	96%	91%	95%	97%	87%	88%	94%	84%	94%
swaptions	94%	94%	97%	82%	103%	102%	90%	93%	97%	77%	80%	80%	96%	89%	103%
vips	96%	93%	95%	82%	103%	101%	95%	91%	93%	87%	80%	78%	97%	79%	85%
x264	94%	94%	96%	82%	103%	102%	106%	85%	88%	98%	96%	98%	101%	89%	98%

Heterogeneous

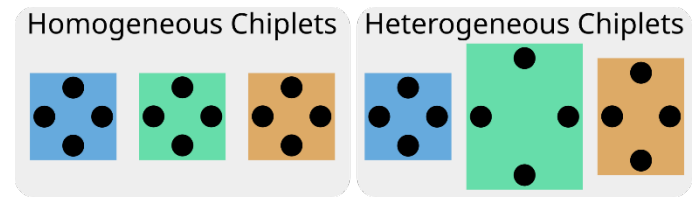
32 cores

	Region 1			Region 2			Region 3			Region 4			Region 5		
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA
blackscholes	87%	89%	87%	92%	97%	92%	103%	100%	97%	100%	101%	108%	93%	98%	94%
bodytrack	84%	87%	84%	92%	97%	91%	145%	111%	126%	110%	106%	107%	89%	93%	89%
canneal	84%	86%	84%	92%	97%	91%	95%	97%	95%	86%	85%	84%	89%	98%	92%
dedup	87%	89%	87%	92%	97%	91%	79%	93%	78%	88%	94%	92%	98%	96%	100%
ferret	87%	89%	87%	92%	97%	91%	95%	102%	98%	98%	99%	99%	89%	87%	87%
fluidanimate	87%	89%	87%	92%	97%	91%	89%	91%	88%	88%	89%	87%	91%	95%	92%
swaptions	84%	87%	84%	92%	97%	91%	83%	88%	82%	99%	96%	101%	85%	91%	86%
vips	87%	90%	87%	92%	97%	91%	98%	101%	99%	91%	87%	94%	88%	82%	95%
x264	84%	87%	84%	92%	97%	92%	104%	95%	100%	102%	106%	98%	87%	92%	87%

64 cores

	Region 1			Region 2			Region 3			Region 4			Region 5		
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA
blackscholes	89%	93%	80%	96%	96%	88%	106%	121%	101%	107%	117%	119%	106%	110%	95%
bodytrack	85%	90%	79%	96%	96%	88%	82%	135%	69%	117%	141%	115%	104%	111%	94%
canneal	85%	90%	79%	96%	96%	88%	101%	106%	96%	98%	92%	89%	107%	108%	93%
dedup	89%	93%	79%	96%	96%	88%	122%	132%	103%	94%	90%	97%	96%	81%	81%
ferret	89%	93%	80%	96%	96%	88%	102%	121%	101%	103%	116%	94%	92%	87%	84%
fluidanimate	89%	93%	79%	96%	96%	88%	93%	93%	91%	95%	89%	87%	105%	106%	93%
swaptions	85%	90%	79%	96%	96%	88%	88%	93%	85%	101%	120%	96%	118%	113%	99%
vips	89%	94%	80%	96%	96%	88%	103%	111%	99%	86%	93%	101%	86%	96%	97%
x264	86%	90%	79%	96%	96%	88%	100%	109%	90%	98%	105%	92%	99%	105%	90%

Evaluation using Partial Traffic Trace Regions



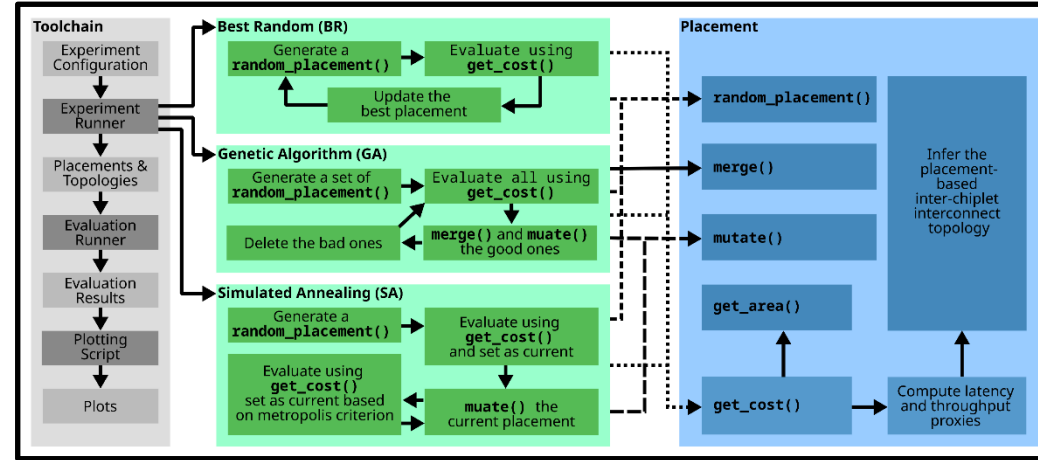
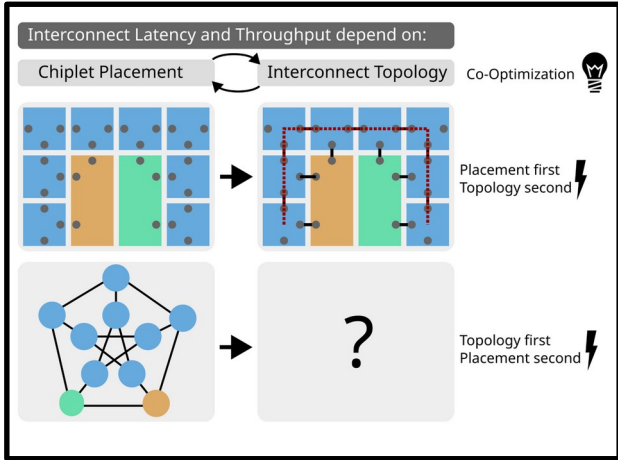
Homogeneous

	32 cores															64 cores															
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	
	Region 1			Region 2			Region 3			Region 4			Region 5			Region 1			Region 2			Region 3			Region 4			Region 5			
blackscholes	73%	80%	76%	49%	65%	66%	66%	76%	79%	90%	77%	77%	74%	95%	96%	82%	79%	75%	74%	75%	75%	78%	80%	79%	80%	73%	80%	80%	80%	80%	79%
bodytrack	72%	81%	75%	49%	65%	66%	52%	70%	72%	59%	78%	76%	68%	85%	84%	83%	81%	75%	74%	75%	75%	81%	99%	94%	72%	74%	66%	70%	68%	66%	
canneal	72%	81%	75%	49%	65%	66%	70%	83%	83%	74%	77%	76%	61%	80%	84%	83%	81%	75%	74%	75%	75%	76%	77%	74%	79%	78%	78%	63%	64%	69%	
dedup	72%	80%	76%	49%	65%	66%	78%	100%	121%	82%	76%	72%	49%	55%	54%	81%	79%	75%	74%	75%	75%	94%	90%	85%	83%	77%	79%	87%	89%	92%	
ferret	73%	80%	76%	49%	65%	66%	67%	93%	86%	65%	82%	80%	81%	69%	71%	81%	79%	75%	74%	75%	75%	78%	79%	72%	72%	72%	71%	77%	78%	77%	
fluidanimate	73%	80%	76%	49%	65%	66%	70%	75%	72%	68%	71%	71%	60%	77%	75%	81%	79%	75%	74%	75%	75%	82%	80%	79%	76%	74%	75%	72%	76%	74%	
swaptions	72%	81%	75%	49%	65%	66%	76%	83%	79%	64%	91%	80%	73%	99%	97%	83%	81%	75%	74%	75%	75%	85%	83%	77%	66%	64%	60%	81%	80%	78%	
x264	73%	81%	75%	49%	65%	66%	54%	66%	65%	64%	82%	82%	68%	91%	89%	82%	81%	75%	74%	75%	75%	71%	73%	68%	76%	81%	79%	78%	79%	77%	

Heterogeneous

	32 cores															64 cores														
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA
	Region 1			Region 2			Region 3			Region 4			Region 5			Region 1			Region 2			Region 3			Region 4			Region 5		
blackscholes	80%	76%	83%	89%	70%	83%	90%	79%	86%	98%	99%	89%	93%	74%	86%	87%	81%	93%	105%	101%	91%	114%	102%	104%	96%	102%	98%	108%	108%	96%
bodytrack	78%	74%	81%	89%	70%	83%	94%	88%	92%	96%	82%	93%	86%	71%	81%	83%	78%	90%	105%	101%	91%	107%	107%	73%	140%	116%	106%	102%	101%	100%
canneal	78%	74%	81%	89%	70%	83%	93%	79%	89%	90%	80%	85%	87%	72%	79%	83%	78%	90%	105%	101%	91%	104%	100%	98%	87%	90%	90%	98%	95%	91%
dedup	80%	76%	83%	89%	70%	83%	78%	60%	71%	87%	87%	88%	102%	84%	85%	87%	81%	93%	105%	101%	91%	144%	111%	114%	88%	90%	91%	93%	84%	85%
ferret	80%	75%	83%	89%	70%	83%	98%	80%	92%	99%	83%	90%	84%	77%	80%	87%	81%	93%	105%	101%	91%	111%	104%	97%	100%	95%	101%	84%	94%	91%
fluidanimate	80%	75%	83%	89%	70%	83%	89%	84%	85%	83%	80%	81%	88%	73%	80%	87%	81%	93%	105%	101%	91%	88%	91%	91%	84%	84%	91%	96%	95%	93%
swaptions	78%	74%	81%	89%	70%	83%	79%	73%	80%	94%	77%	85%	82%	67%	77%	83%	78%	90%	105%	101%	91%	89%	85%	94%	107%	109%	105%	113%	105%	108%
x264	77%	74%	81%	89%	70%	83%	83%	73%	77%	88%	85%	93%	85%	71%	79%	83%	78%	91%	105%	101%	91%	101%	93%	90%	95%	93%	95%	98%	97%	91%

Conclusion



More of SPCL's research:

youtube.com/@sycl **180+ Talks**

twitter.com/sycl_eth **1.4K+ Followers**

github.com/sycl **4K+ Stars**

... or sycl.ethz.ch

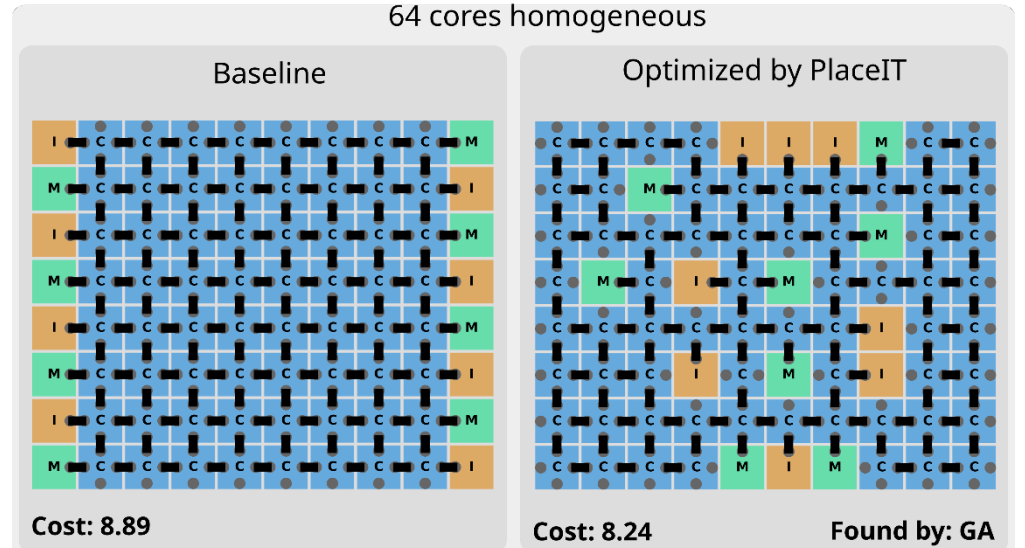
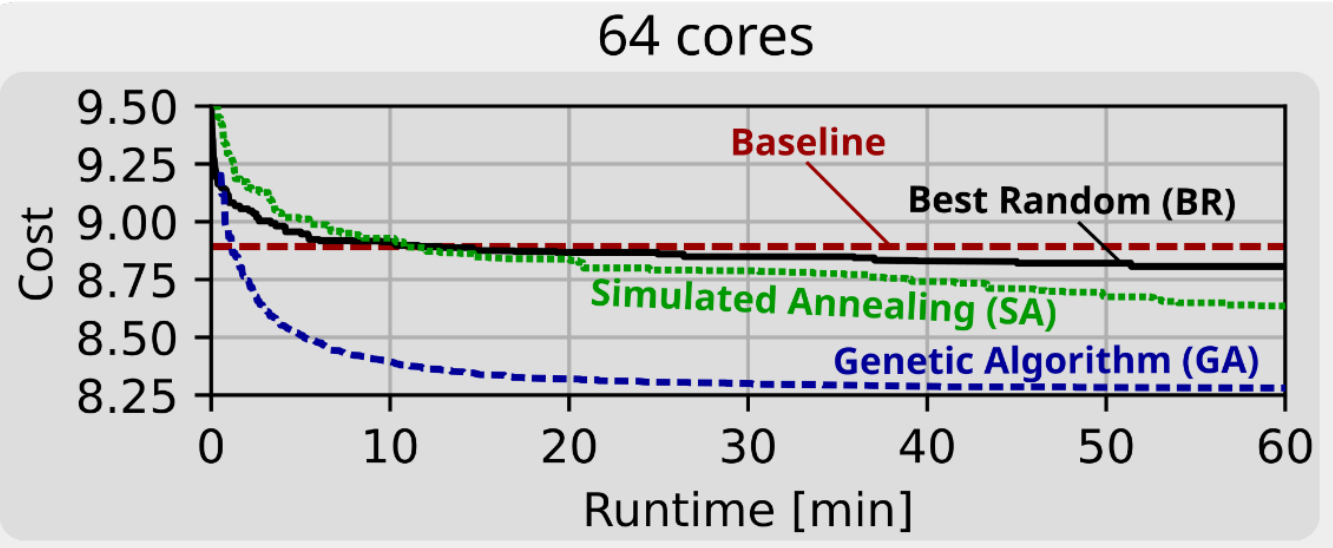
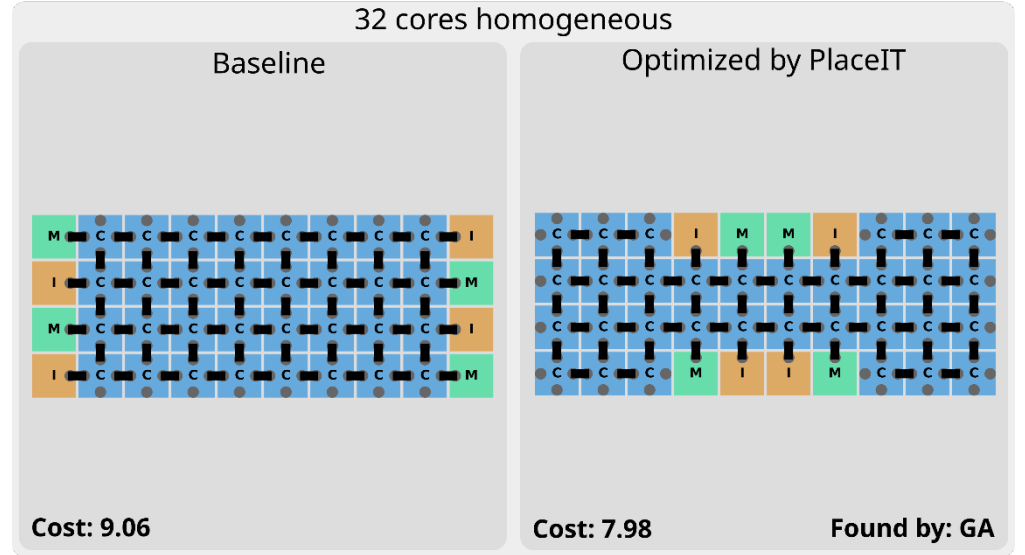
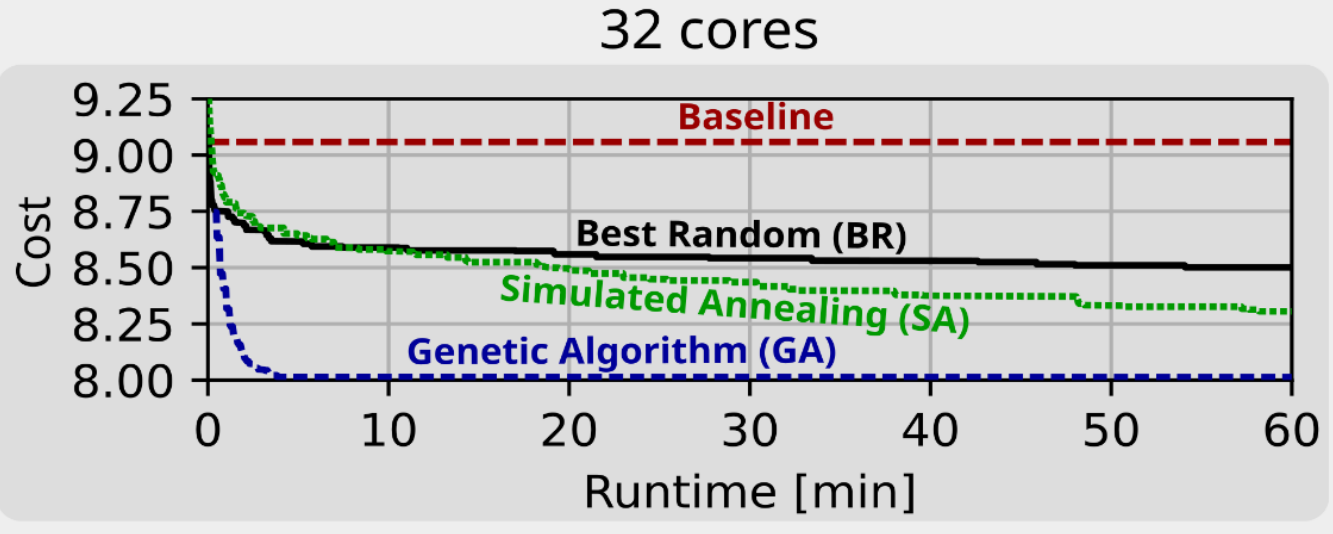


	Homogeneous Chiplets										Heterogeneous Chiplets																			
	32 cores					64 cores					32 cores					64 cores														
blackscholes	92%	88%	88%	77%	79%	83%	93%	85%	82%	95%	96%	99%	109%	101%	98%	96%	93%	80%	82%	88%	91%	100%	95%	91%	95%	93%	94%			
bodytrack	90%	86%	89%	77%	79%	81%	76%	75%	69%	88%	81%	81%	101%	90%	92%	94%	94%	97%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
cannal	90%	86%	89%	77%	79%	83%	98%	90%	89%	95%	85%	88%	104%	87%	87%	87%	93%	93%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
dedup	92%	86%	88%	77%	79%	83%	119%	122%	87%	86%	92%	75%	81%	87%	87%	93%	93%	93%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
ferret	92%	86%	89%	77%	79%	83%	102%	94%	94%	95%	92%	88%	87%	74%	86%	96%	93%	93%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
fluidanimate	92%	86%	89%	77%	79%	83%	88%	82%	85%	85%	78%	82%	93%	82%	83%	96%	93%	93%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
swaptions	90%	86%	89%	77%	79%	83%	90%	91%	91%	84%	85%	81%	109%	101%	102%	94%	94%	94%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
vips	92%	86%	89%	77%	79%	83%	100%	92%	91%	80%	75%	78%	78%	68%	74%	96%	93%	93%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
x264	90%	86%	89%	77%	79%	83%	81%	71%	71%	87%	84%	91%	108%	98%	94%	94%	94%	94%	82%	103%	102%	94%	88%	92%	100%	95%	95%	89%	78%	87%
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA
	Region 1			Region 2			Region 3			Region 4			Region 5			Region 1			Region 2			Region 3			Region 4			Region 5		

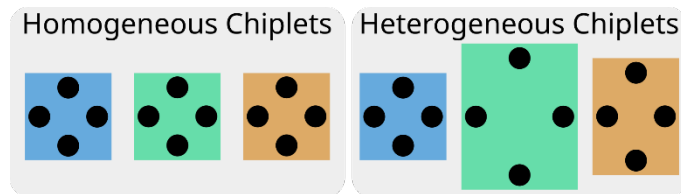
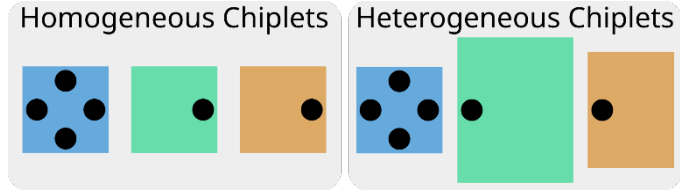
Backup Slides

Optimizing Chips with Homogeneously Shaped Chiplets

■ Compute-Chiplet
 ■ Memory-Chiplet
 ■ IO-Chiplet
 ● PHY
 — D2D Link



Evaluation using Synthetic Traces



Latency relative to the baseline (lower is better)

Traffic Type	32 cores homog.			64 cores homog.			32 cores heterog.			64 cores heterog.		
C2C	124%	123%	127%	112%	118%	117%	107%	111%	102%	108%	110%	116%
C2M	89%	86%	90%	95%	91%	92%	86%	87%	83%	93%	90%	91%
C2I	91%	84%	88%	99%	90%	96%	94%	90%	90%	94%	93%	93%
M2I	69%	50%	55%	88%	71%	79%	87%	81%	82%	92%	85%	81%
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA

Latency relative to the baseline (lower is better)

Traffic Type	32 cores homog.			64 cores homog.			32 cores heterog.			64 cores heterog.		
C2C	125%	123%	123%	117%	118%	116%	109%	112%	115%	112%	113%	116%
C2M	76%	72%	75%	79%	79%	78%	85%	82%	83%	86%	89%	92%
C2I	76%	72%	74%	80%	78%	81%	82%	79%	82%	86%	83%	86%
M2I	38%	41%	41%	65%	53%	60%	60%	48%	51%	73%	66%	81%
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA

Throughput relative to the baseline (higher is better)

Traffic Type	32 cores homog.			64 cores homog.			32 cores heterog.			64 cores heterog.		
C2C	55%	57%	50%	62%	49%	57%	68%	87%	94%	47%	49%	51%
C2M	100%	94%	83%	93%	107%	70%	93%	98%	90%	105%	105%	122%
C2I	106%	93%	74%	84%	91%	104%	90%	102%	100%	96%	71%	92%
M2I	95%	98%	81%	98%	108%	101%	91%	101%	98%	151%	149%	151%
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA

Throughput relative to the baseline (higher is better)

Traffic Type	32 cores homog.			64 cores homog.			32 cores heterog.			64 cores heterog.		
C2C	91%	82%	96%	56%	81%	80%	108%	109%	104%	59%	57%	55%
C2M	128%	163%	135%	146%	137%	161%	131%	130%	116%	104%	131%	120%
C2I	154%	152%	150%	153%	151%	124%	131%	106%	86%	114%	121%	110%
M2I	105%	109%	101%	100%	104%	82%	126%	96%	118%	188%	178%	175%
	BR	GA	SA	BR	GA	SA	BR	GA	SA	BR	GA	SA