香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

ArchPower: Dataset for Architecture-Level Power Modeling of Modern CPU Design

Qijun Zhang, Yao Lu, Mengming Li, Shang Liu, Zhiyao Xie

Hong Kong University of Science and Technology

{qzhangcs, yludf, mengming.li, sliudx}@connect.ust.hk, eezhiyao@ust.hk

Outline

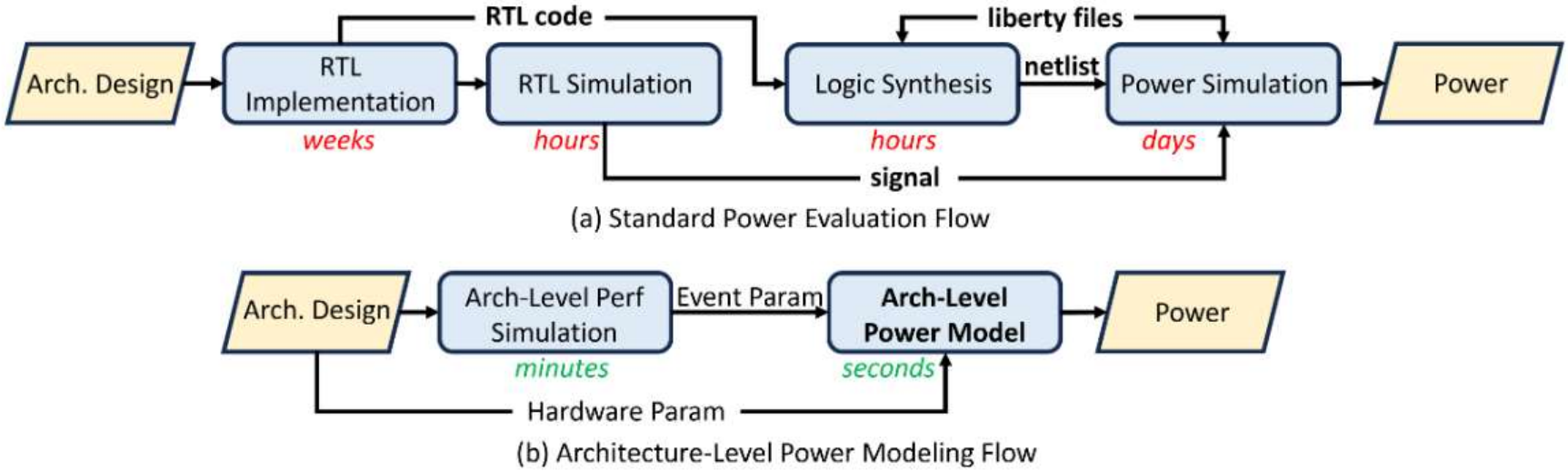
1 Introduction

2 ArchPower Dataset

3 Evaluation

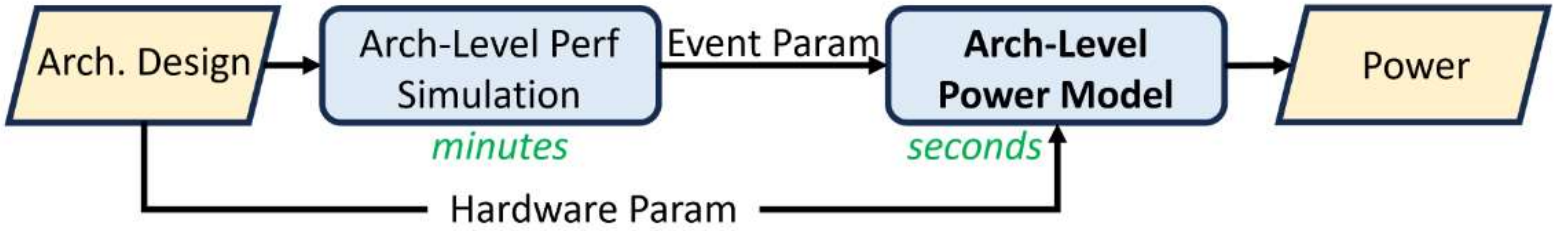
Architecture-Level Power Model

- **Power efficiency** is a critical design objective in microprocessor design
- A high demand for **fast**, yet **high-fidelity** architecture-level power modeling



Architecture-Level Power Model

- Input:
  - **Hardware parameters**, e.g. FetchWidth, DecodeWidth, DCacheWays
  - **Event parameters**, e.g. the number of DCache Miss, Branch Misprediction
- Output:
  - **Power**



Lack of Open-Source Dataset

香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

Problem 1:

- There is ***no open-source dataset*** ML-based architecture-level power models
  - Existing works built their solutions on in-house datasets

Problem 2:

- In-house datasets also have ***other limitations***:
  - ***Limitation 1:*** Do not include **SRAM** in their implementation
  - ***Limitation 2:*** Do not adopt the **clock-gating** technique
  - ***Limitation 3:*** Only collected based on a **single** CPU architecture

Lack of Open-Source Dataset

Problem 1:

- There is ***no open-source dataset*** ML-based architecture-level power models
  - Existing

We propose ***ArchPower***, the first **open-source** dataset for ML-based architecture-level power models

Pr

- In-he
  - ***Limitation 1:*** Do
  - ***Limitation 2:*** Do not adopt the **clock-gating** technique
  - ***Limitation 3:*** Only collected based on a **single** CPU architecture

Outline

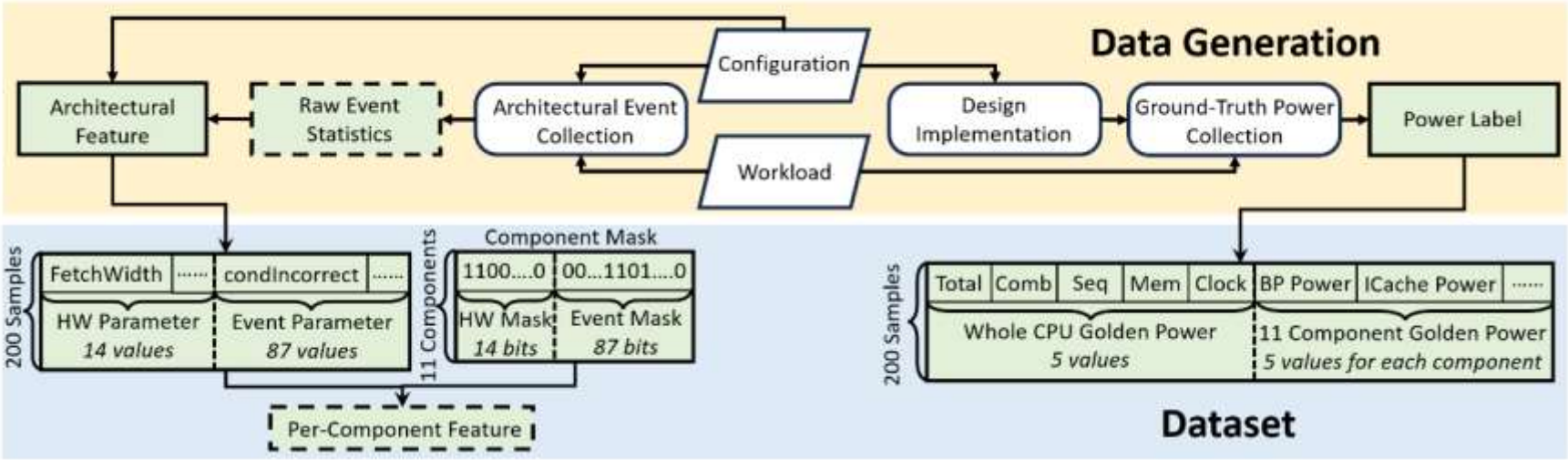
1 Introduction

2 **ArchPower Dataset**

3 Evaluation

Dataset Overview

- Consists of 25 × 8 = 200 data samples
  - 25 CPU configurations
  - 8 different workloads
- Providing both architecture-level ***features*** and power ***labels***



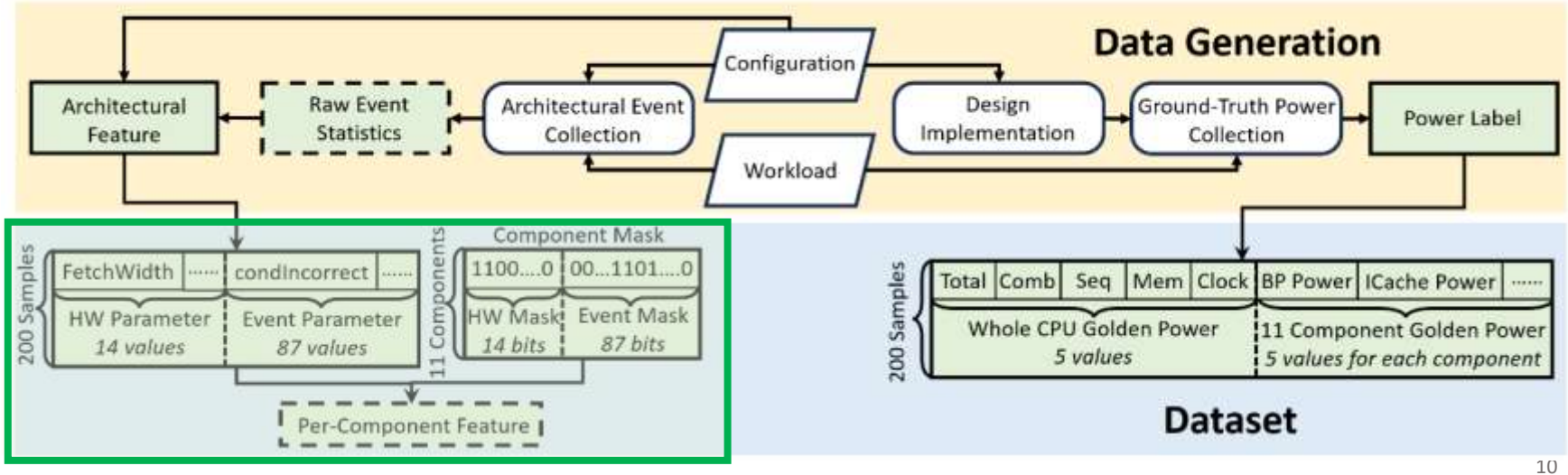
CPU Configurations

- **15 BOOM** configurations (B1-B15) and **10 XiangShan** configurations (X1-X10)
- Some **representative** CPU configurations:

Hardware Parameter	B1	B2	B4	B6	B7	B9	B11	B13	B15	X1	X3	X5	X7	X8	X10
FetchWidth	4	4	4	8	8	8	8	8	8	4	4	4	8	8	8
DecodeWidth	1	1	2	2	3	3	4	5	5	2	2	3	4	4	5
FetchBufferEntry	5	8	8	24	18	30	32	30	40	8	24	24	24	32	24
RobEntry	16	32	64	80	81	114	128	125	140	16	48	64	81	96	112
IntPhyRegister	36	53	64	88	88	112	128	108	140	36	68	80	88	110	108
FpPhyRegister	36	48	56	72	88	112	128	108	140	36	68	80	88	110	108
LDQ/STQEntry	4	8	12	20	16	32	32	24	36	16	24	24	24	32	32
BranchCount	6	8	10	14	14	16	20	18	20	7	7	7	7	7	7
Mem/FpIssueWidth	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
IntIssueWidth	1	1	1	2	2	3	4	5	5	2	2	4	4	6	6
DCache/ICacheWay	2	4	4	8	8	8	8	8	8	4	8	4	8	8	8
DTLBEntry	8	8	8	16	16	32	32	32	32	8	16	8	16	16	32
MSHREntry	2	2	2	4	4	4	4	8	8	2	4	2	4	4	4
ICacheFetchBytes	2	2	2	4	4	4	4	4	4	2	2	2	2	2	2

Architectural Power Modeling Feature

- Each sample has 14 + 87 = 101 features
  - 14: hardware parameters
  - 87: event parameters
- Provides a component mask to select the features for each component



Architectural Power Modeling Feature

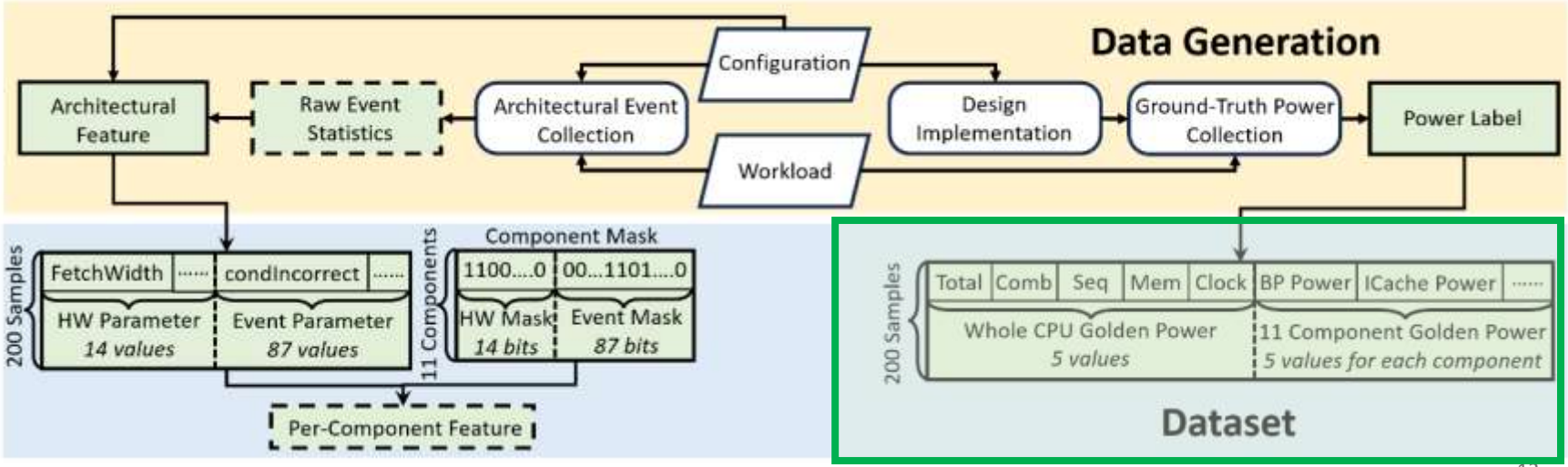
香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Hardware parameters and event parameters of each component:

Component	Hardware parameters of each component	Event parameters of each component
BP	FetchWidth, BranchCount	BTBLookups, condPredicted, condIncorrect, commit.branches
IFU	FetchWidth, DecodeWidth FetchBufferEntry, ICacheFetchBytes	fetch.{insts, branches, cycles}, numRefs, numStoreInsts, numInsts, decode.{runCycles, blockedCycles, decodedInsts}, numBranches, intInstQueueReads, intInstQueueWrites, intInstQueueWakeupAccesses, fpInstQueueReads, fpInstQueueWrites, fpInstQueueWakeupAccesses
ICache	ICacheWay, ICacheFetchBytes	overallAccesses, overallMisses, ReadReq.mshrHits, ReadReq.mshrMisses, tagAccesses
RNU	DecodeWidth	intLookups, renamedOperands, fpLookups, renamedInsts, runCycles, blockCycles, committedMaps
ROB	DecodeWidth, RobEntry	reads, writes
ISU	DecodeWidth, Mem/FpIssueWidth, IntIssueWidth	IssuedMemRead, IssuedMemWrite, IssuedFloatMemRead, IssuedFloatMemWrite, IssuedIntAlu, IssuedIntMult, IssuedIntDiv, IssuedFloatMult, IssuedFloatDiv
Regfile	DecodeWidth, IntPhyRegister, FpPhyRegister	intRegfileReads, fpRegfileReads, intRegfileWrites, fpRegfileWrites, functionCalls
FU Pool	Mem/FpIssueWidth, IntIssueWidth	intAluAccesses, fpAluAccesses
LSU	LDQ/STQEntry, MemIssueWidth	MemRead, InstPrefetch, MemWrite
DCache	DCacheWay, DCacheTLBEntry, DCacheMSHR, MemIssueWidth	ReadReq.accesses, WriteReq.accesses, ReadReq.misses, tagAccesses, WriteReq.misses, overallMisses, MshrHits, MshrMisses
CPU Level	-	totalIpc, totalCpi, numCycles, idleCycles, numLoadInsts, numSquashedInsts, committedInsts, commit.{numDist::mean, memRefs}, mmu.dtb.{accesses, misses}, iew.writebackCount, numIssuedDist::mean, statIssuedInstType_0::total, fuBusy, mmu.itb.{accesses, misses}, conflictingLoads, conflictingStores, insertedLoads, insertedStores, mem_ctrls.{readReqs, writeReqs, bytesReadSys}, icache.tags.totalRefs, dcache.{overallAccesses::total, tags.totalRefs}

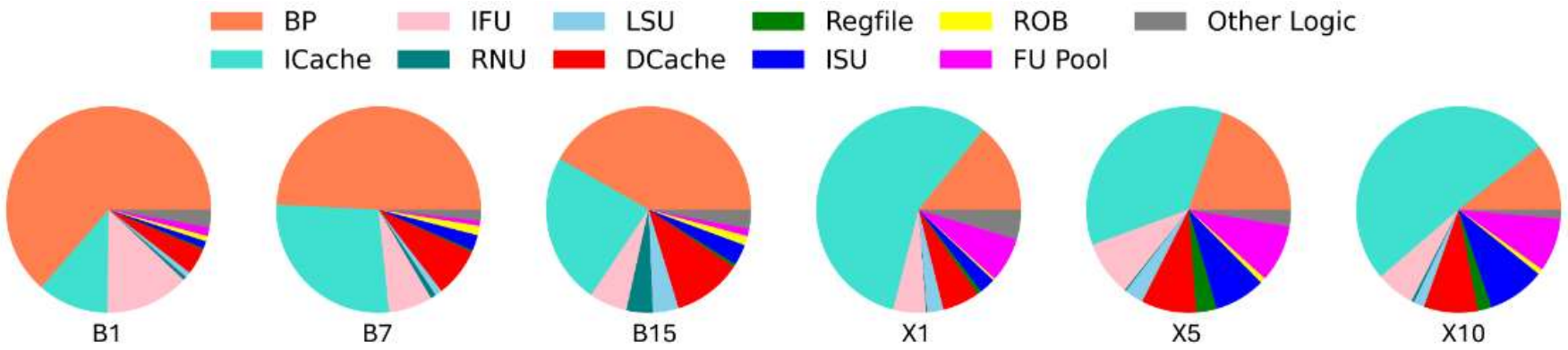
Power Label

- Each sample has (1 + 11) × (1 + 4) = 60 **fine-grained** power labels
  - 1 + 11: The whole CPU and 11 components
  - 1 + 4: Further decouple the power into four power groups
    - combinational logic power, sequential logic power, memory power, and clock power



Power Label

- The power distributions across 11 components of 6 different CPU configurations with different scales:



Outline

1 Introduction

2 ArchPower Dataset

3 **Evaluation**

Benchmarked Models

香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

- Evaluate **six** architecture-level power models
- **Two** analytical models:
  - (a) McPAT
  - (b) McPAT-Plus
- **Four** ML-based models:
  - (c) McPAT-Calib
  - (d) McPAT-Calib-Component
  - (e) McPAT-Calib-CompGroup
  - (f) PANDA

Training-Testing Data Setup

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Set up **three** training-testing scenarios
- 1) **Balance**
  - **Evenly** select configurations as training configurations based on the scale
  - B1, B8, and B15 for BOOM, X1, X6, and X10 for XiangShan
- 2) **Small**
  - Select the **smallest** configurations as available training configurations
  - B1, B2, and B3 for BOOM, X1, X2, and X3 for XiangShan
- 3) **Large**
  - Select the **largest** configurations as available training configurations
  - B13, B14, and B15 for BOOM, X8, X9, and X10 for XiangShan

Power Prediction Accuracy

- Accuracy comparison between our selected six models

Scenario	McPAT				McPAT-Plus				McPAT-Calib			
	BOOM		XiangShan		BOOM		XiangShan		BOOM		XiangShan	
	MAPE	R	MAPE	R	MAPE	R	MAPE	R	MAPE	R	MAPE	R
Balance	>100	0.83	>100	0.85	18.1	0.83	29.6	0.85	8.2	0.98	33.2	0.73
Small	>100	0.74	>100	0.77	31.0	0.74	21.6	0.77	34.3	0.76	41.5	0.48
Large	>100	0.83	>100	0.78	28.2	0.83	28.3	0.78	50.6	0.23	90.0	0.14
Average	>100	0.80	>100	0.80	25.8	0.80	26.5	0.80	31.0	0.66	54.9	0.45
Scenario	McPAT-Calib-Component				McPAT-Calib-CompGroup				PANDA			
	BOOM		XiangShan		BOOM		XiangShan		BOOM		XiangShan	
	MAPE	R	MAPE	R	MAPE	R	MAPE	R	MAPE	R	MAPE	R
Balance	6.2	0.98	**14.0**	0.97	**6.2**	**0.98**	15.0	**0.97**	6.8	0.97	19.4	0.9
Small	34.9	0.75	35.4	0.72	35.3	0.75	36.0	0.72	**29.2**	**0.93**	**23.9**	**0.86**
Large	48.9	0.4	81.4	0.31	49.2	0.4	80.5	0.35	**10.4**	**0.98**	**26.3**	**0.82**
Average	30.0	0.71	43.6	0.67	30.2	0.71	43.8	0.68	**15.5**	**0.96**	**23.2**	**0.86**

Conclusion

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

- **ArchPower**, the **first open-source** dataset for ML-based architecture-level power models
- Includes **200 data samples** collected from 25 CPU configurations and 8 workloads
- Consider the **clock-gating** and integrate realistic **SRAM macros** for power label collection

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

THANK YOU!

Qijun Zhang, Yao Lu, Mengming Li, Shang Liu, Zhiyao Xie

Hong Kong University of Science and Technology

{qzhangcs, yludf, mengming.li, sliudx}@connect.ust.hk, eezhiyao@ust.hk