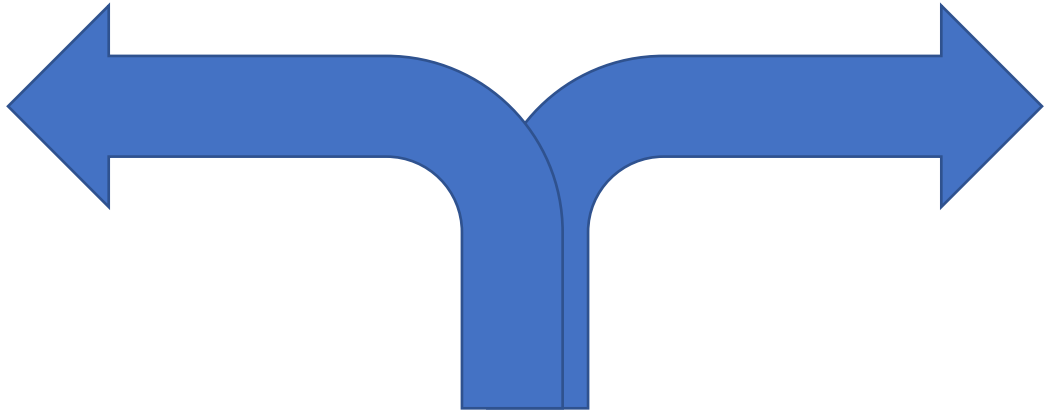
Where do research ideas come from?

Margaret Martonosi
H. T. Adams ‘35 Professor of Computer Science
Princeton University

Where do *successful* research ideas come from?

Margaret Martonosi
H. T. Adams ‘35 Professor of Computer Science
Princeton University

Sorting through ideas…



About the idea itself

- What is the problem?
- Why is it important?
- What have others done about it?
- What can I/we do that is different?
- What special skills, viewpoints, resources or data can we bring to it?

About the idea as a work opportunity

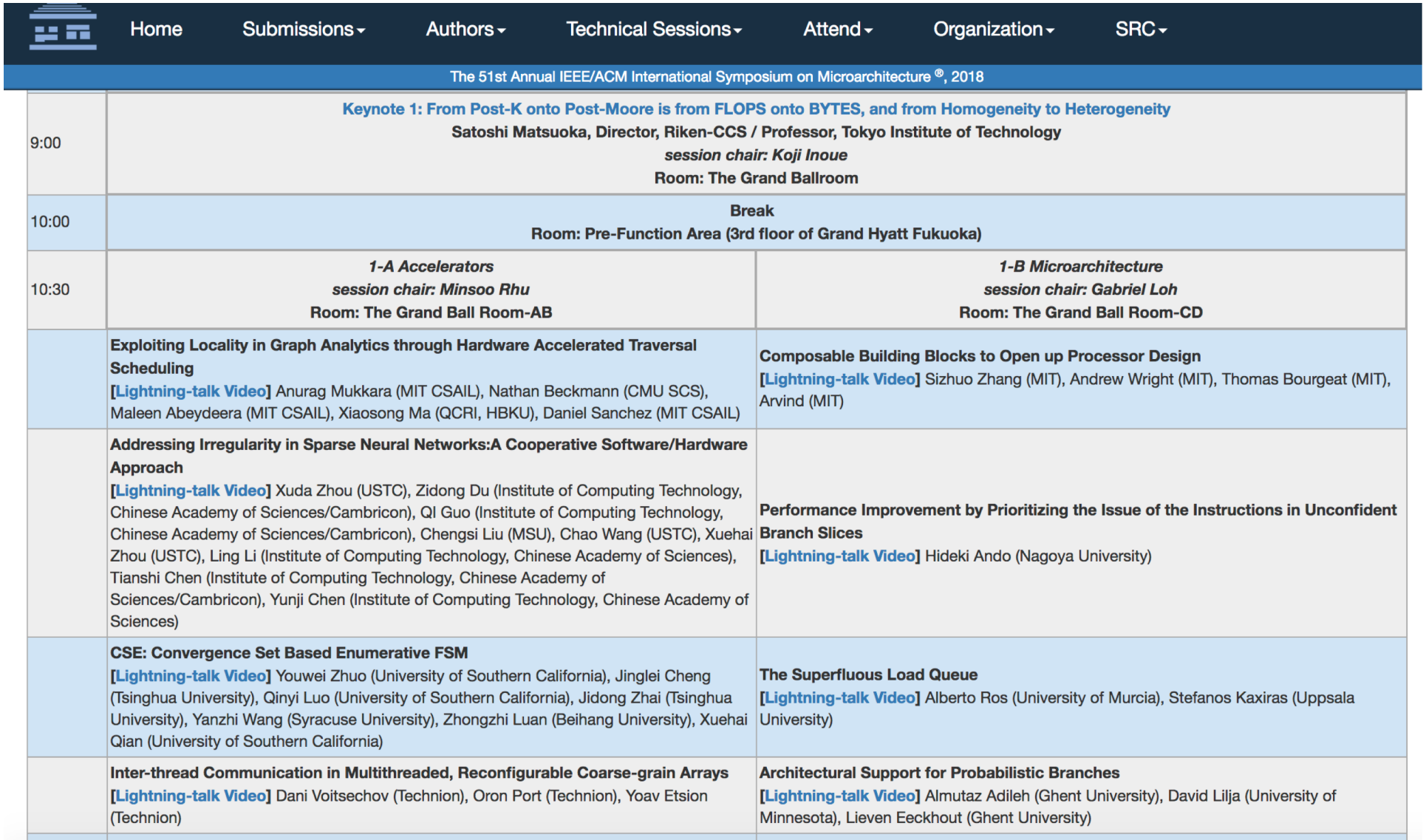
- What will it take to solve it, and do I have some (not all!) of what it takes to solve it?
  - Be open to acquiring new skills.
- Does the team seem fun?
- Will they be respectful of me and my ideas?
- Can I [enjoyably] learn new things by working on this problem with this group of collaborators?

What to do with good ideas?

- Grad school:
  - Can only work on 1-2 at a time yourself
  - For the rest:
    - Suggest to groupmates
    - Work with undergrads
    - Save for when you're less busy if they'll "keep"
- Professor:
  - Find a student to work on them!
  - And/or, Write a proposal!

Keep a notebook or Google Keep or Trello list

0. Reading papers and Attending talks

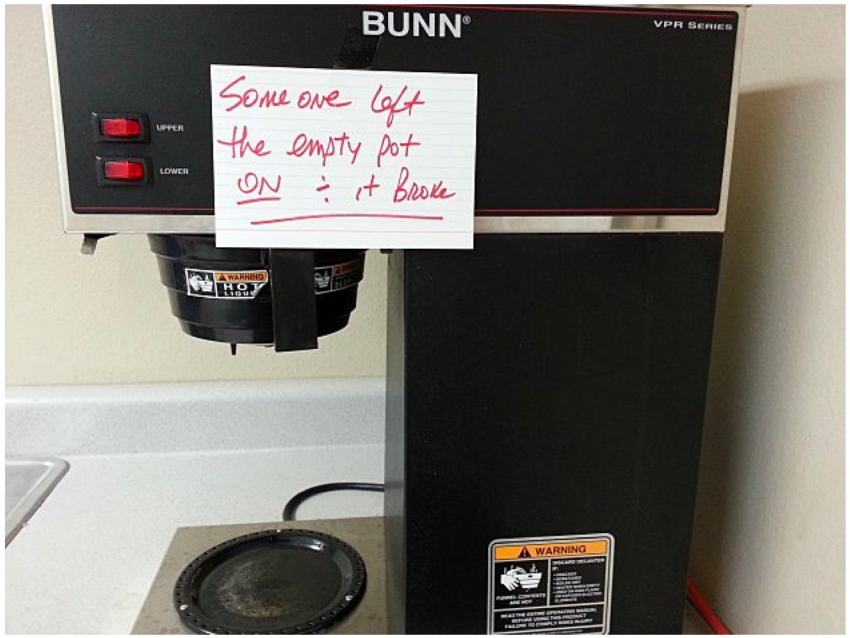


Home | Submissions | Authors | Technical Sessions | Attend | Organization | SRC

The 51st Annual IEEE/ACM International Symposium on Microarchitecture ®, 2018

9:00	**Keynote 1: From Post-K onto Post-Moore is from FLOPS onto BYTES, and from Homogeneity to Heterogeneity** **Satoshi Matsuoka, Director, Riken-CCS / Professor, Tokyo Institute of Technology** ***session chair: Koji Inoue*** **Room: The Grand Ballroom**	
10:00	**Break** **Room: Pre-Function Area (3rd floor of Grand Hyatt Fukuoka)**	
10:30	***1-A Accelerators*** ***session chair: Minsoo Rhu*** **Room: The Grand Ball Room-AB**	***1-B Microarchitecture*** ***session chair: Gabriel Loh*** **Room: The Grand Ball Room-CD**
	Exploiting Locality in Graph Analytics through Hardware Accelerated Traversal Scheduling **[Lightning-talk Video]** Anurag Mukkara (MIT CSAIL), Nathan Beckmann (CMU SCS), Maleen Abeydeera (MIT CSAIL), Xiaosong Ma (QCRI, HBKU), Daniel Sanchez (MIT CSAIL)	**Composable Building Blocks to Open up Processor Design** **[Lightning-talk Video]** Sizhuo Zhang (MIT), Andrew Wright (MIT), Thomas Bourgeat (MIT), Arvind (MIT)
	Addressing Irregularity in Sparse Neural Networks:A Cooperative Software/Hardware Approach **[Lightning-talk Video]** Xuda Zhou (USTC), Zidong Du (Institute of Computing Technology, Chinese Academy of Sciences/Cambricon), Qi Guo (Institute of Computing Technology, Chinese Academy of Sciences/Cambricon), Chengsi Liu (MSU), Chao Wang (USTC), Xuehai Zhou (USTC), Ling Li (Institute of Computing Technology, Chinese Academy of Sciences), Tianshi Chen (Institute of Computing Technology, Chinese Academy of Sciences/Cambricon), Yunji Chen (Institute of Computing Technology, Chinese Academy of Sciences)	**Performance Improvement by Prioritizing the Issue of the Instructions in Unconfident Branch Slices** **[Lightning-talk Video]** Hideki Ando (Nagoya University)
	CSE: Convergence Set Based Enumerative FSM **[Lightning-talk Video]** Youwei Zhuo (University of Southern California), Jinglei Cheng (Tsinghua University), Qinyi Luo (University of Southern California), Jidong Zhai (Tsinghua University), Yanzhi Wang (Syracuse University), Zhongzhi Luan (Beihang University), Xuehai Qian (University of Southern California)	**The Superfluous Load Queue** **[Lightning-talk Video]** Alberto Ros (University of Murcia), Stefanos Kaxiras (Uppsala University)
	Inter-thread Communication in Multithreaded, Reconfigurable Coarse-grain Arrays **[Lightning-talk Video]** Dani Voitsechov (Technion), Oron Port (Technion), Yoav Etsion (Technion)	**Architectural Support for Probabilistic Branches** **[Lightning-talk Video]** Almutaz Adileh (Ghent University), David Lilja (University of Minnesota), Lieven Eeckhout (Ghent University)

1. Friendly conversations



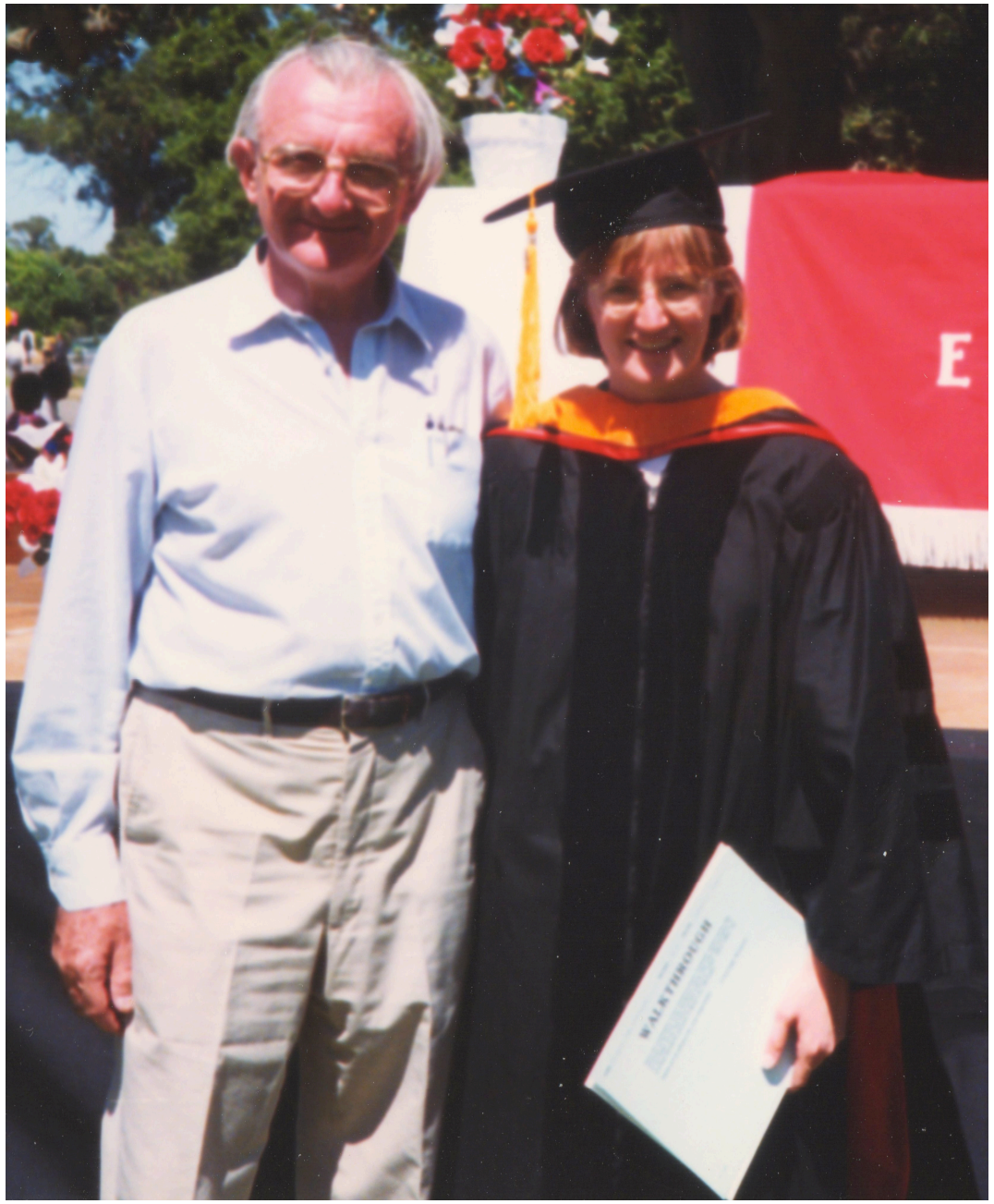
- Not this: “Hello you’ve never met me but what are your best research ideas that we can collaborate on?”
- Better:
  - “What are you working on right now?”
  - “What do you think are some key trends that will shape our field 5-10 years from now?”
  - “I’ve always wondered – why do you think we don’t see more papers about <___> at this conference?”
  - And as you know someone better “Hey I was working on <___> today and I thought of your work because of <___>…”

- Example: I started QC research 12 years ago when a device physicist at the department coffee machine started talking about his idea for a new qubit.

2. Thesis adviser + …

- Listen to your adviser: if you've got a reasonably good adviser, they have good ideas and they have more experience than you.
- + Add to your adviser's good ideas: With your own wisdom, from papers you read and from colleagues you talk to.

This is not my thesis adviser →



3. A desire to “break out of your box”

ZebraNet Project Accomplishments

Goal

Energy-efficient detailed wildlife tracking in disconnected regions, to support studies of wildlife ecology, land management, and protection of endangered species.

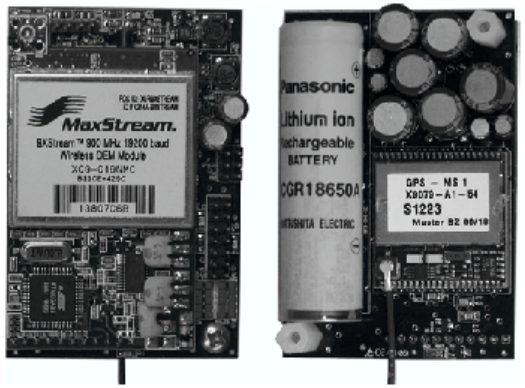
Results

Complete hardware, software, package designs:

- 2 deployments on zebras in Kenya
- ~5 distinct generations, following Moore's Law improvements in processor and memory + Improvements in GPS
- Over-the-air SW updates for remotely-deployed sensor nodes.

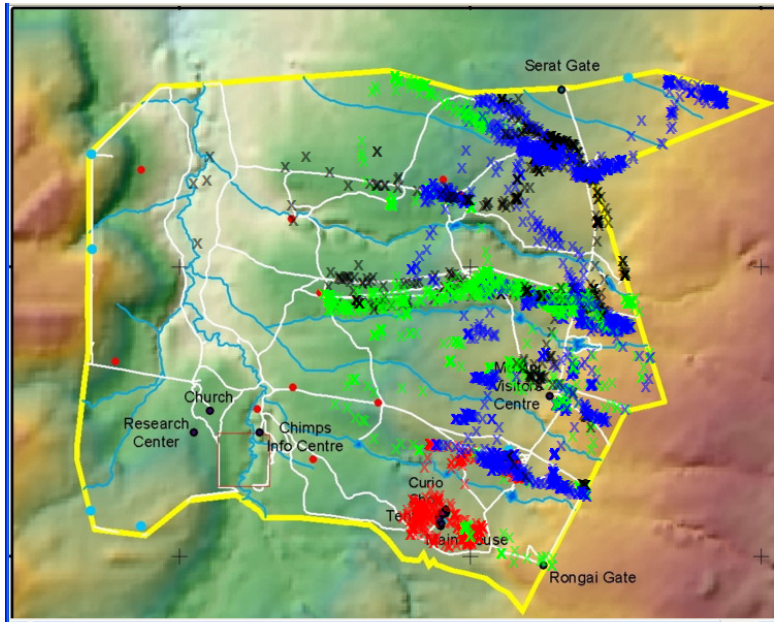
First-ever Mobile Sensing & Novel DTN Protocols

- Delay-tolerant networks for energy-efficient, peer-to-peer data collection in disconnected regions.



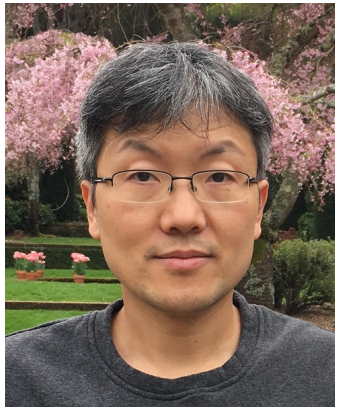
Then ➞ Now: Longer-Term Impact

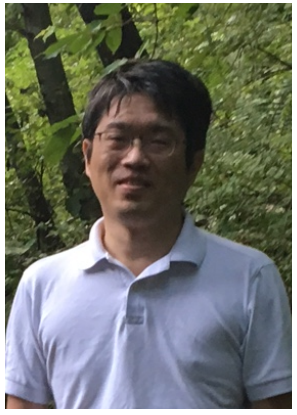
- Broad Applications of Delay-Tolerant Networking
  - For rural, disconnected communities, for energy-efficiency, and even for data privacy…
  - CarTel, DakNet, C-LINK,
- Human mobility & Opportunistic Mobile sensing
  - Collars -> Cellphones – track humans!
  - Energy-efficient protocols
  - Mobile sensing -> Urban planning
  - More info in ACM GetMOBILE retrospective: http://mrmgroup.cs.princeton.edu/papers/p14-martonosi.pdf
- Inspired higher-functionality commercial radio collars and tracking techniques:
  - Computer vision algorithms to automatically identify tracked individuals.
  - Rich sensor technology
  - Regional aggregation via cell towers. Mobile cellular data communication
- Social networks analysis in Wildlife Ecology
  - Interactions between tracked individuals
  - Citizen science tracking and crowdsourced data collection
  - IBEIS and the Hotspotter algorithms now called WILDBOOK (see IBEIS.org).



Princeton ZebraNet 2002-2007

- **Students**: Chris Sadler, Pei Zhang, Ting Liu, Ilya Fischoff, Philo Juang, Yong Wang, Hidekazu Oki
- **Faculty**: Margaret Martonosi, Daniel Rubenstein, Steve Lyon, Li-Shiuan Peh
- **Thanks to**: National Science Foundation + Princeton University Mpala Research Centre



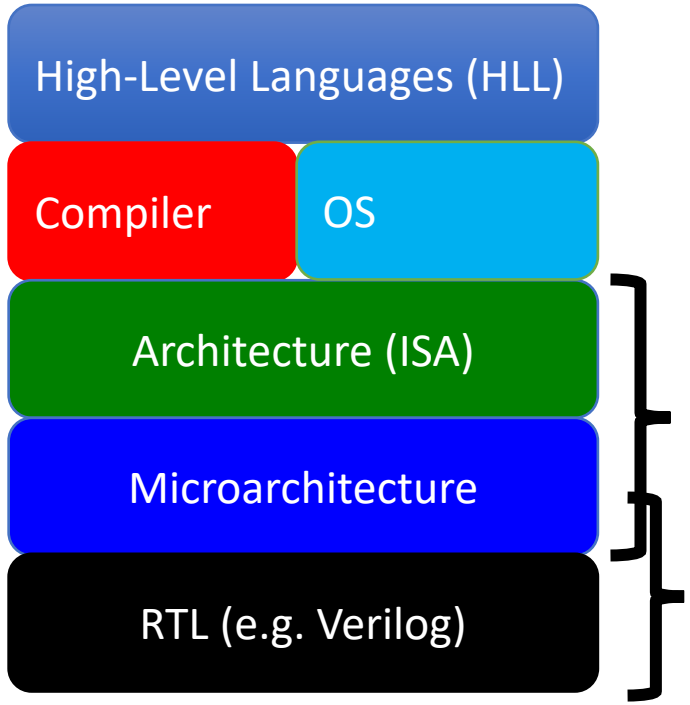


4. Listening to Smart Students

- Dan Lustig: “I really think there are important research problems related to memory consistency model verification.”

- MRM:

The Check Suite: An Ecosystem of Tools For Verifying Memory Consistency Model Implementations



CheckMate [Micro '18]

PipeProof [Micro '18] [Nominated for Best Paper Award]

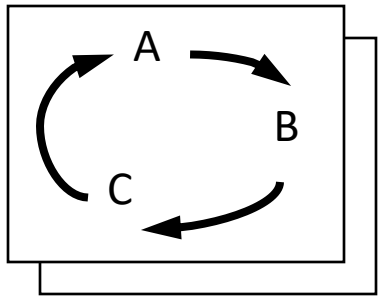
TriCheck [ASPLOS '17] [IEEE MICRO Top Picks]

COATCheck [ASPLOS '16] [IEEE MICRO Top Picks]

PipeCheck [Micro '14] [IEEE MICRO Top Picks]

CCICheck [Micro '15] [Nominated for Best Paper Award]

RTLCheck [Micro '17] [IEEE MICRO Top Picks Honorable Mention]



Our Approach

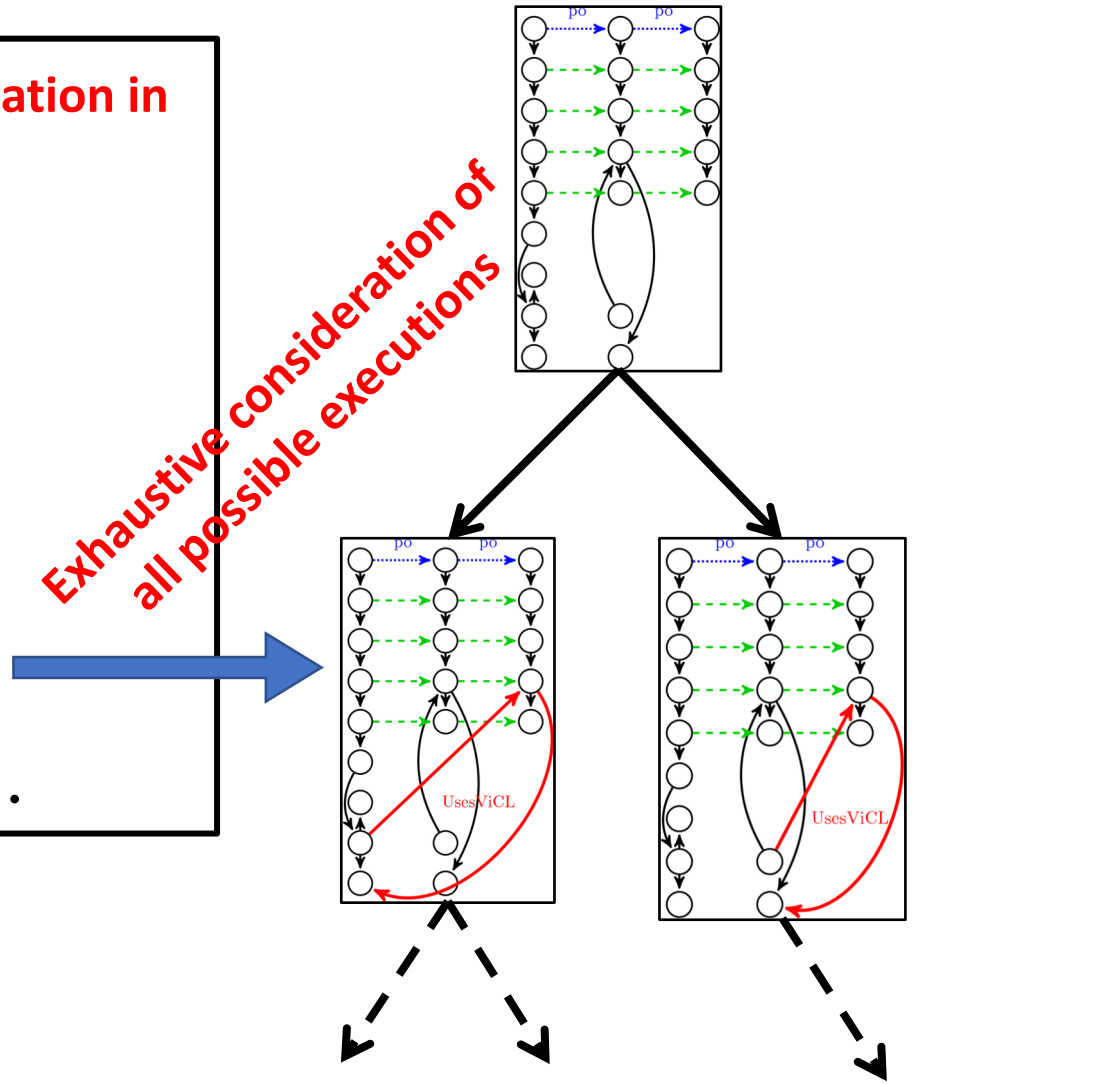
- Axiomatic specifications -> Happens-before graphs
- Check Happens-Before Graphs via Efficient SMT solvers
  - Cyclic => A->B->C->A… Can't happen
  - Acyclic => Scenario is observable

Check: Formal, Axiomatic Models and Interfaces

```
Axiom "PO_Fetch":
forall microops "i1",
forall microops "i2",
SameCore i1 i2 /\ ProgramOrder i1 i2 =>
  AddEdge ((i1, Fetch), (i2, Fetch), "PO").

Axiom "Execute_stage_is_in_order":
forall microops "i1",
forall microops "i2",
SameCore i1 i2 /\
 EdgeExists ((i1, Fetch), (i2, Fetch)) =>
   AddEdge ((i1, Execute), (i2, Execute), "PPO").
```

Microarchitecture Specification in *μSpec* DSL



Microarchitectural happens-before (μhb) graphs

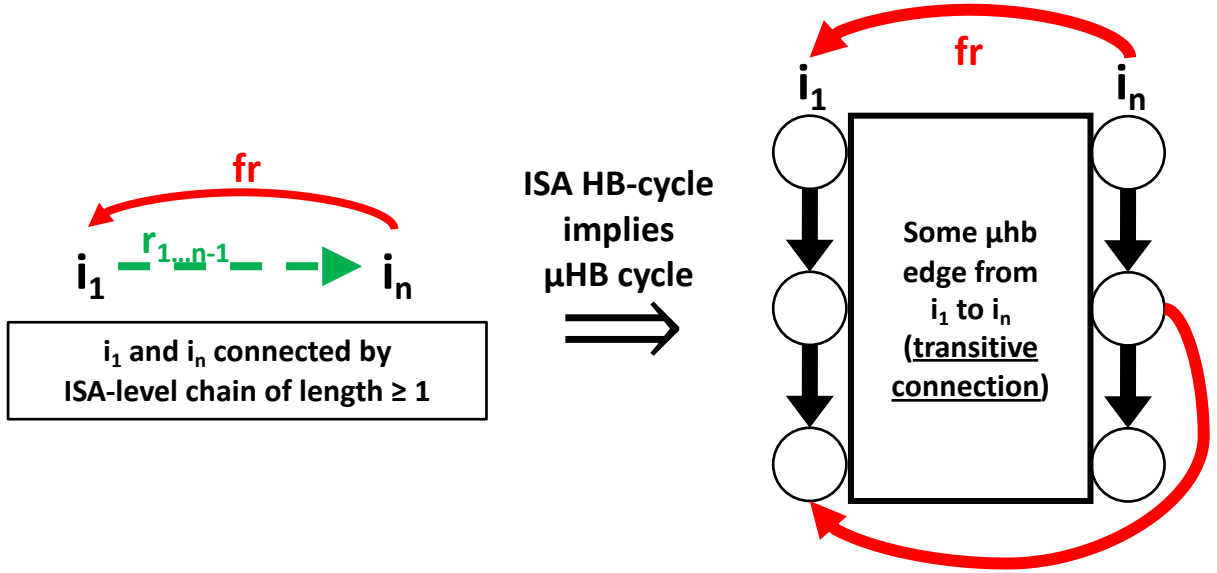
4. Listening to Smart Students

- Yatin Manerkar: “I want people to to stop asking us about litmus tests.”

- MRM:

PipeProof: From Litmus Tests to All Programs

- Litmus Tests are bounded and may be incomplete
  - May not cover all MCM requirements
- MCM concisely & fully specified via ISA-level relations
  - Eg: SC is $acyclic(po \cup co \cup rf \cup fr)$
- How to prove μArch implements MCM for arbitrary programs of unbounded length?

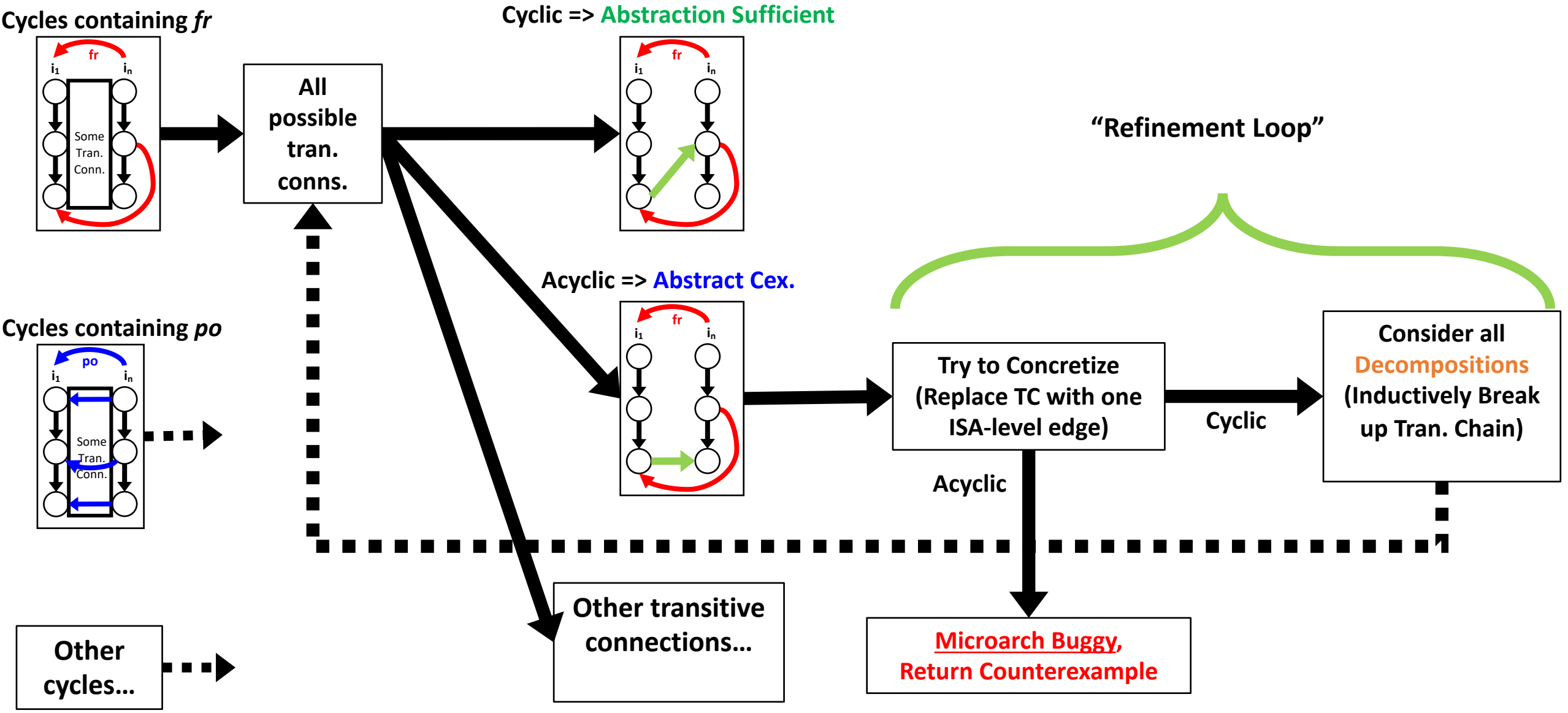


Must construct an **inductive** proof that ISA-level cycle ⟹ μHB cycle

Nominated for MICRO-51 Best Paper Award

[Manerkar, Lustig, Martonosi, Gupta. MICRO-51. October, 2018. http://check.cs.princeton.edu/papers/manerkar_MICRO18.pdf]

PipeProof Proof Sketch



4. Listening to Smart Students

- Caroline Trippel: “I think we can link memory models and security issues.”

- MRM:

Attack Discovery & Synthesis: What We Would Like

1. Specify system to study	Formal interface and specification of given system implementation
2. Specify attack pattern	E.g. Subtle event sequences during program's execution
3. Synthesis	Either output synthesized attacks. Or determine that none are possible

Attack Discovery & Synthesis:
CheckMate TL;DR

1. Specify system to study
2. Specify attack pattern
3. Synthesis

- **What we did**: Developed a tool to do this, based on the uHB graphs from previous sections.
- **Results**: Automatically synthesized Spectre and Meltdown, as well as two new distinct exploits and many variants.

[Trippel, Lustig, Martonosi. https://arxiv.org/abs/1802.03802]
[Trippel, Lustig, Martonosi. MICRO 2018. October, 2018] http://check.cs.princeton.edu/papers/ctrippel_MICRO51.pdf

5. Proving the Naysayers Wrong

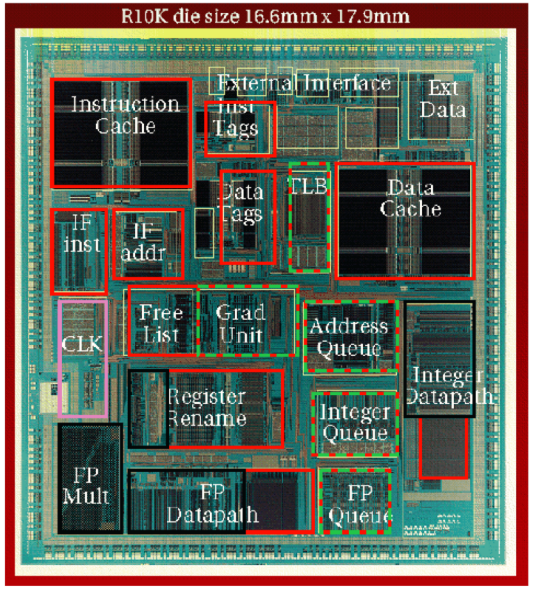
Told to me in 1999…

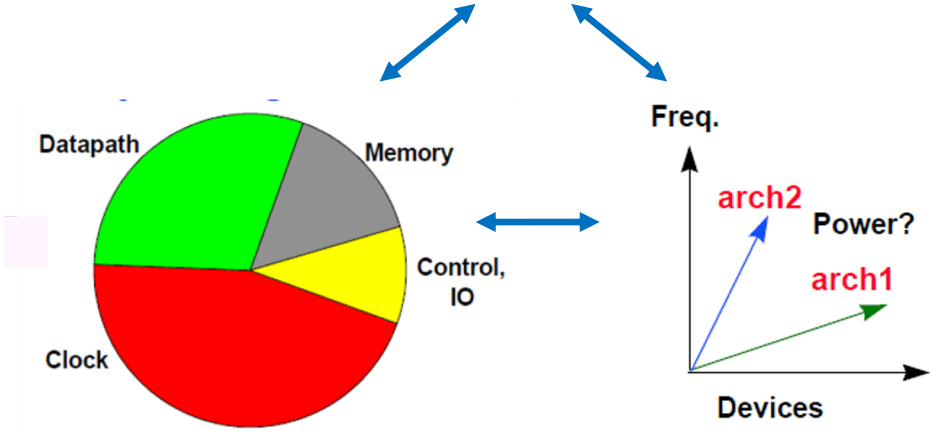
“Power is for the circuits/devices people. Architects shouldn’t do power.”

→ Some of the time, you need to do what people tell you not to do.

Wattch: Context and Retrospective

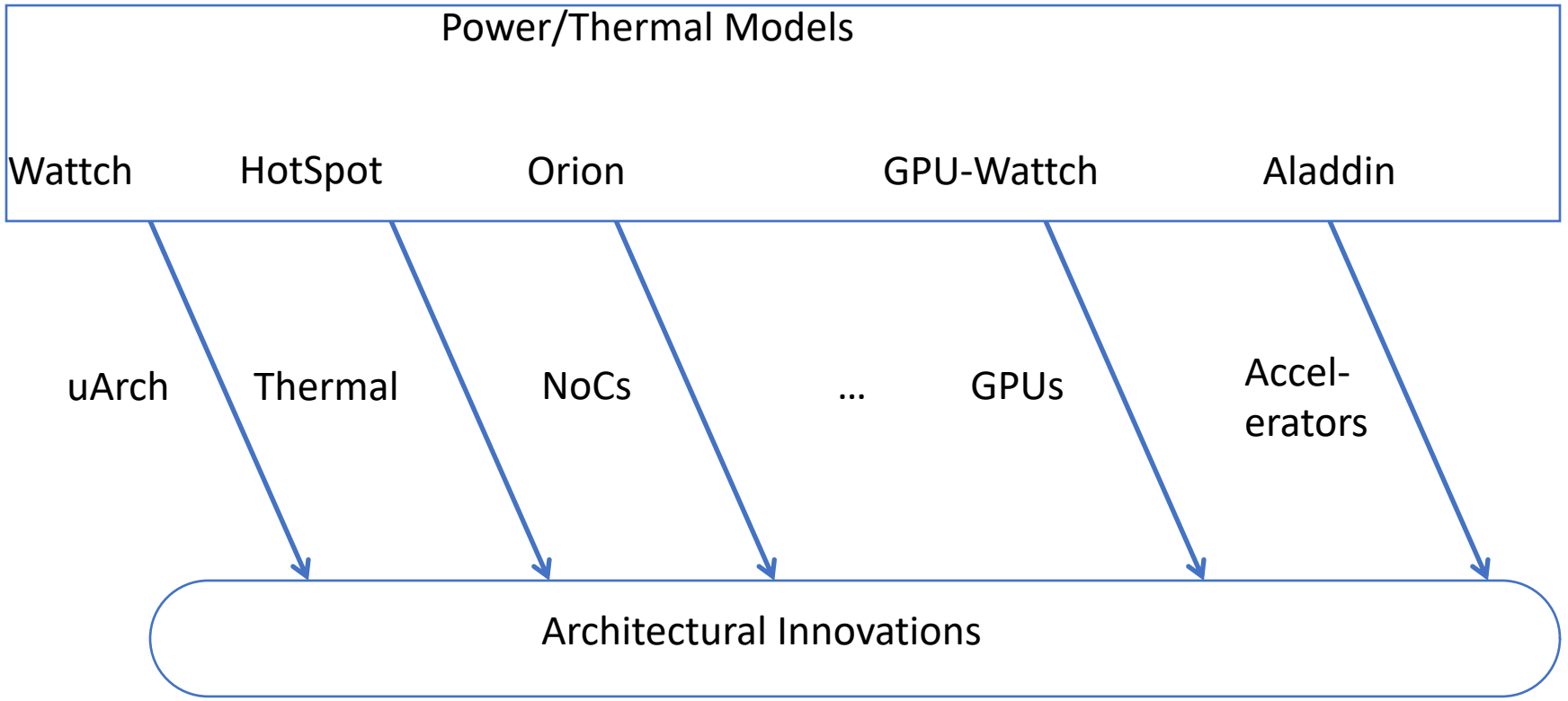
- Early 1990's: Circuit, device engineers handle physical design issues like power. Not yet architects.
- Growing sense that early-stage (pre-RTL) architecture techniques had good power leverage.
- Late 1990's: Power-aware architectural ideas; speculation control, narrow bitwidth, low-power caches
- Where does the power go in a CPU? How to quantify the impact of arch-level power opportunities?





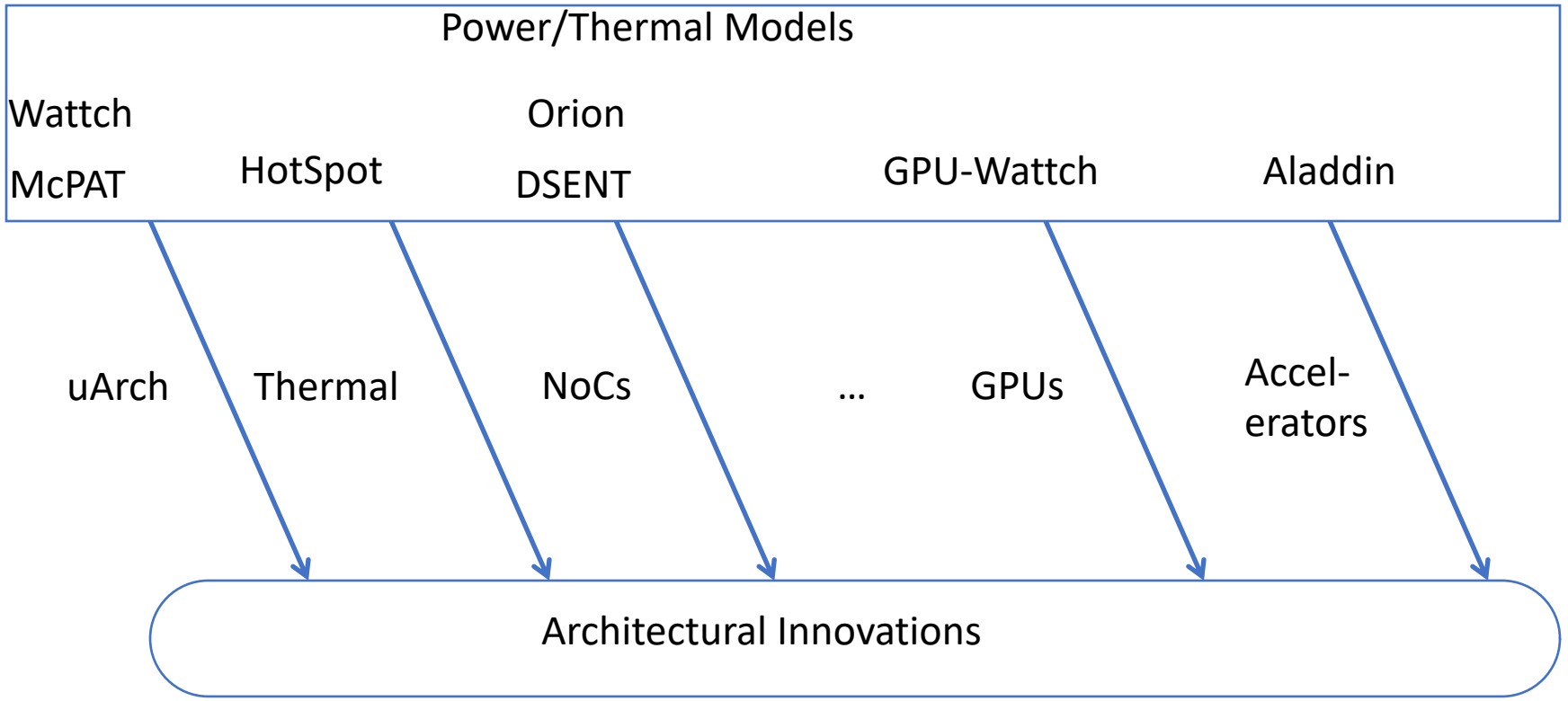
[Brooks, Tiwari, Martonosi. ISCA 2000. (2015 ISCA Influential Paper Award)

Wattch: Impacts

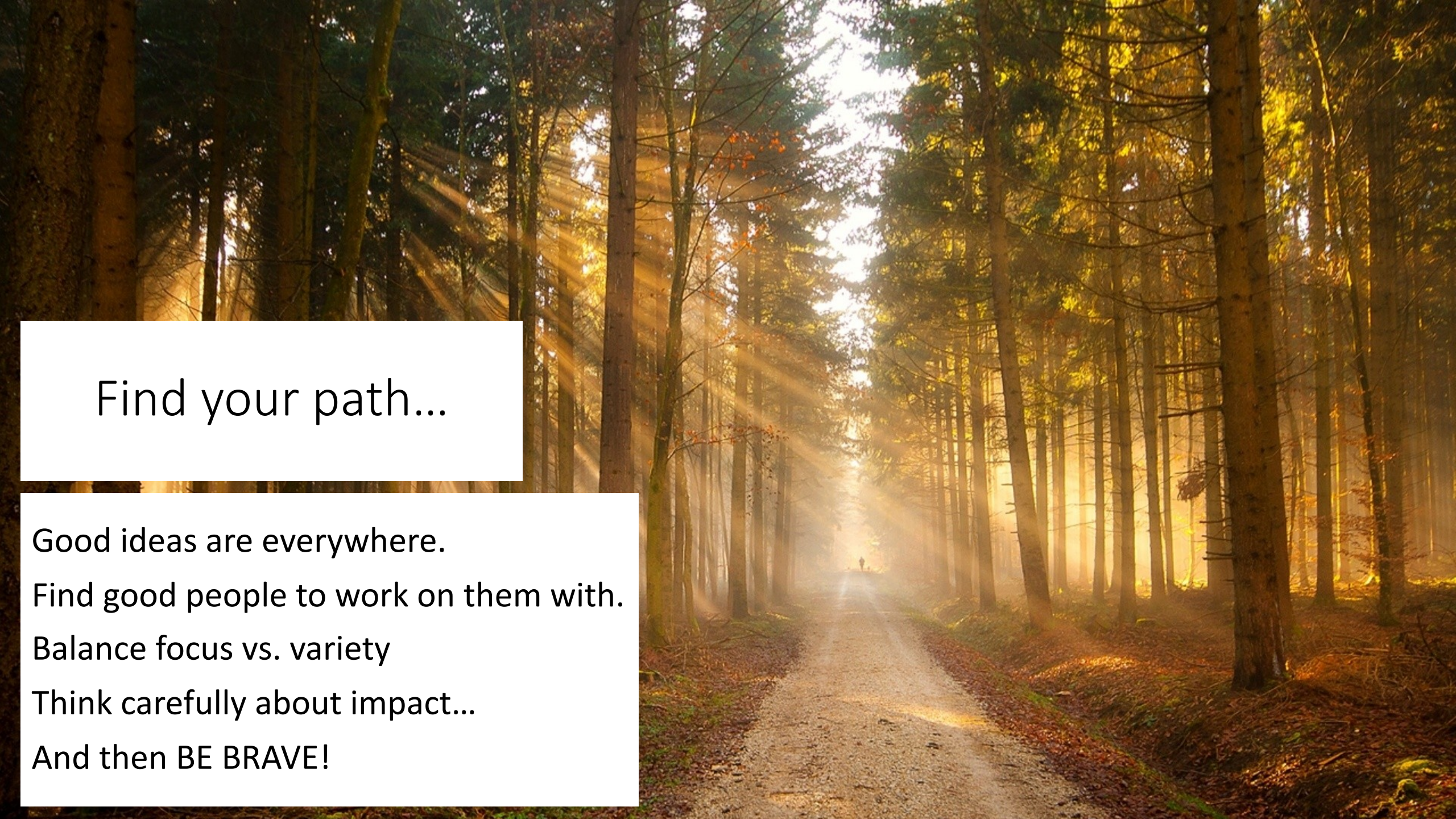


- Wattch: Pre-RTL power models could be useful and could fuel architectural innovation…

Wattch: Impacts



- From ~2000-present: Power transitions from being largely unstudied by architects to being a primary metric alongside performance



Find your path…

Good ideas are everywhere.

Find good people to work on them with.

Balance focus vs. variety

Think carefully about impact…

And then BE BRAVE!