Architecting for Intermittence

Joshua San Miguel
University of Wisconsin-Madison
jsanmiguel@wisc.edu

Arm Research Summit 2019

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Everything is Computing…

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

…Computing is Everything

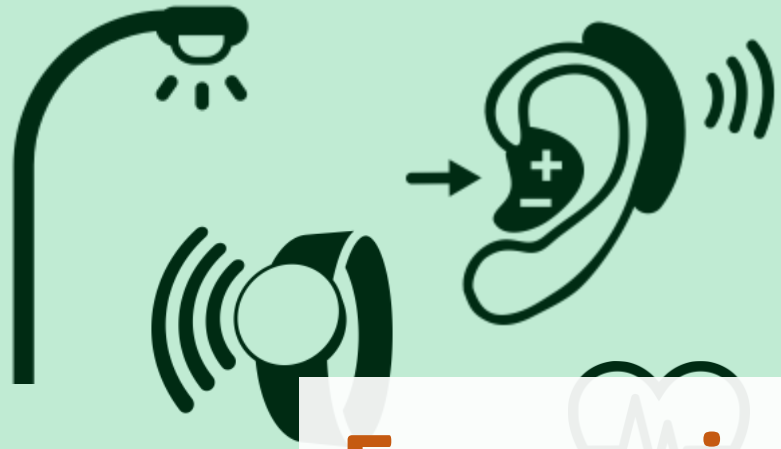
long lifetime

quick response

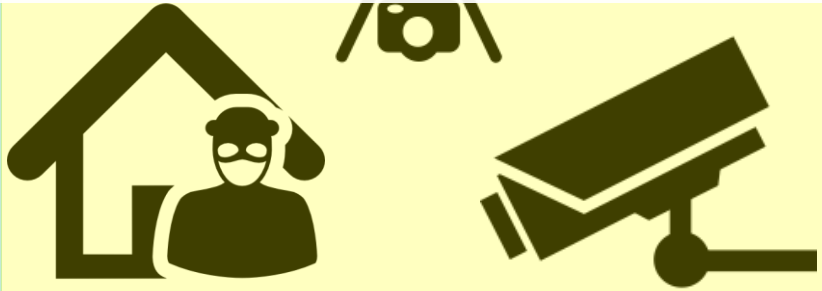
advanced capability

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

…Computing is Everything



Energy is the common denominator



long lifetime

quick response

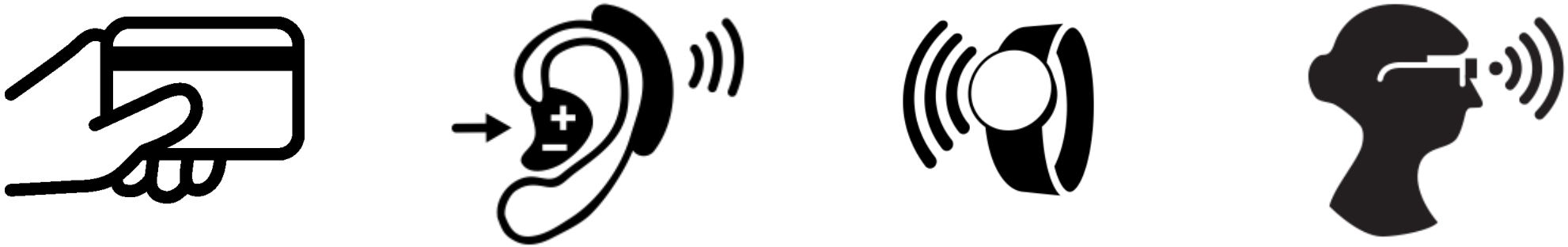
advanced capability

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

Energy-Harvesting Devices

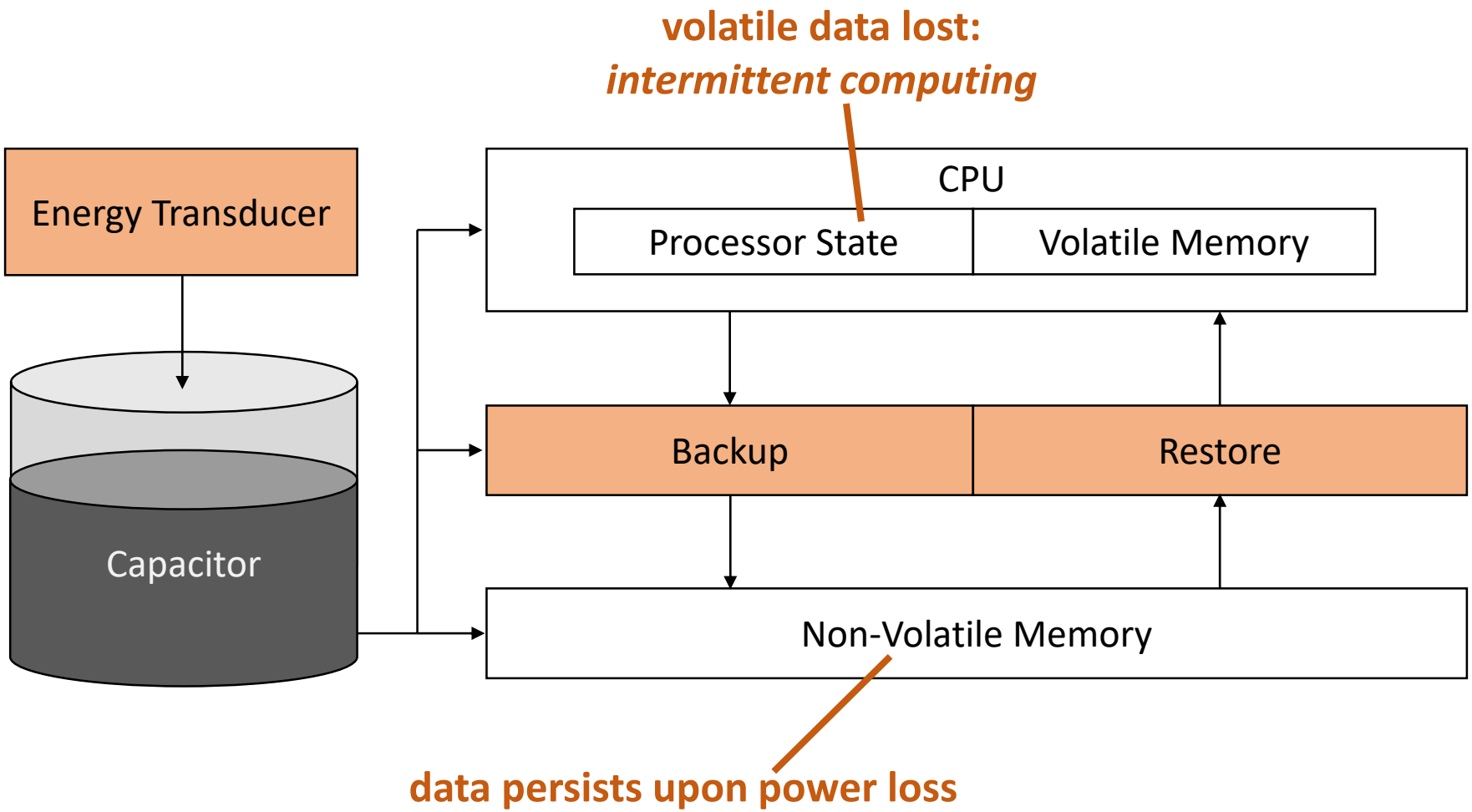
RF
Motion (Human)
Light (Indoor)
Thermal (Human)
Motion (Mechanical)
Thermal (Industrial)
Light (Outdoor)

µW
mW
W



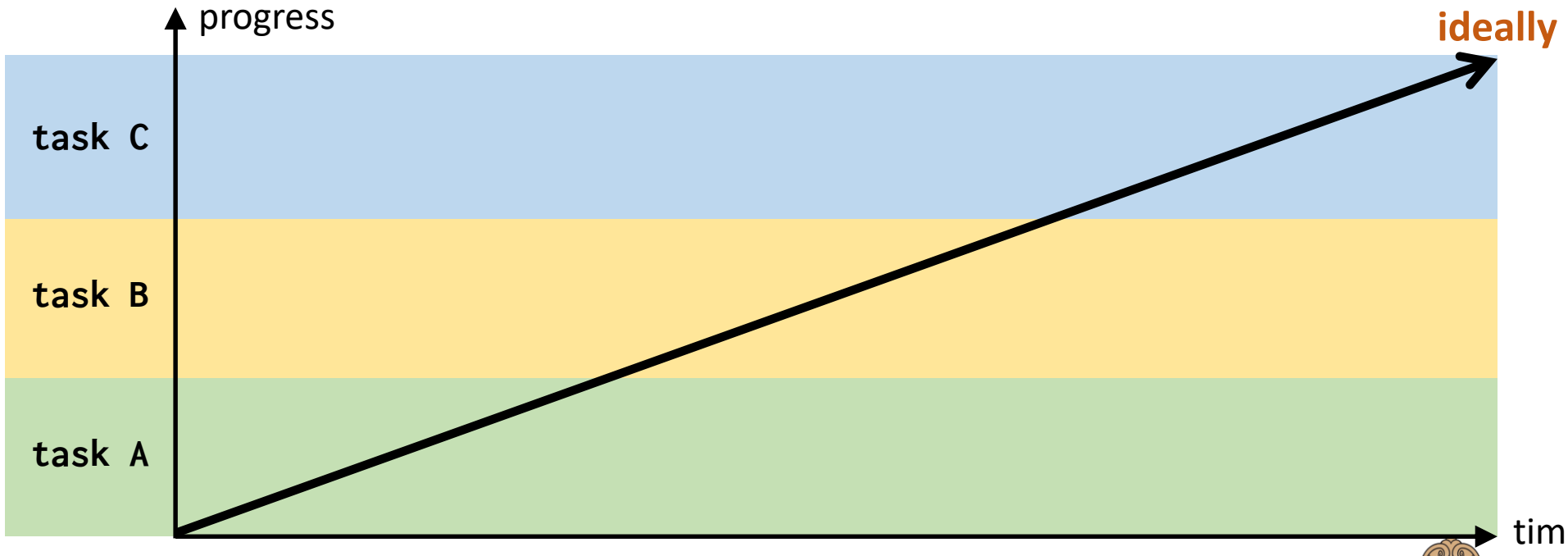
Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Intermittent Computing



Intermittent Computing

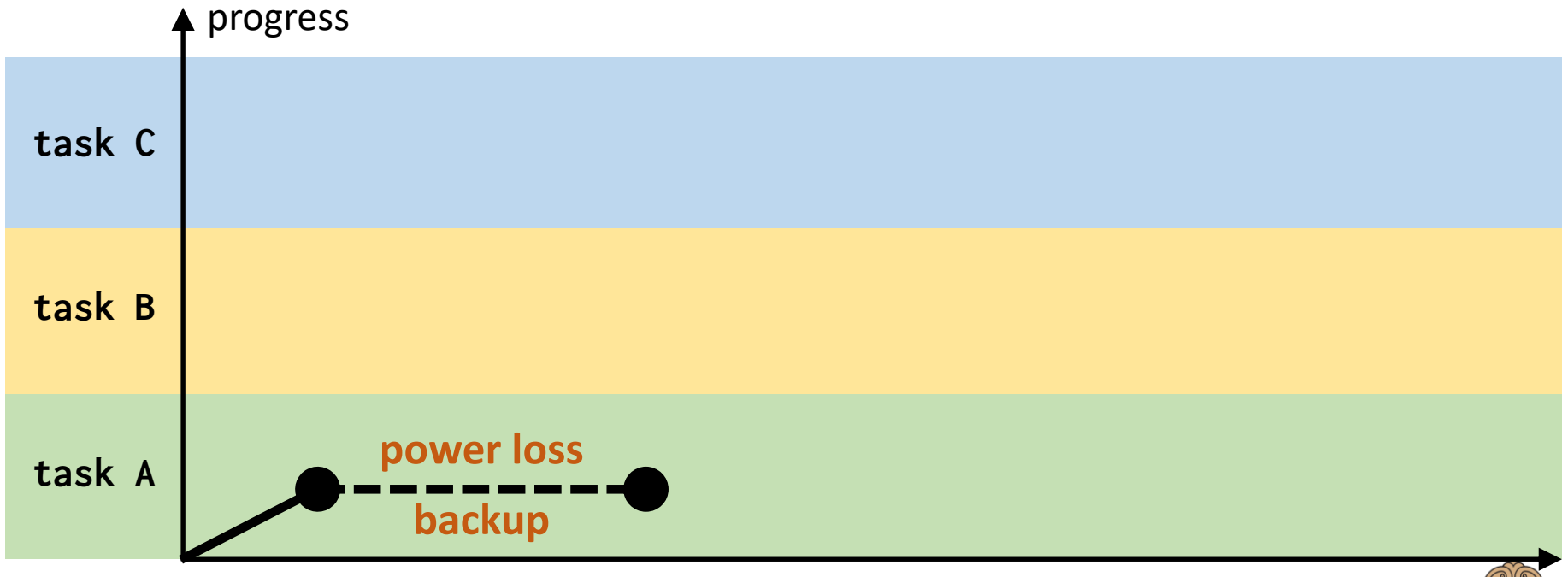
Computation may stop at any point in the program and cannot resume until the device has harvested sufficient energy



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

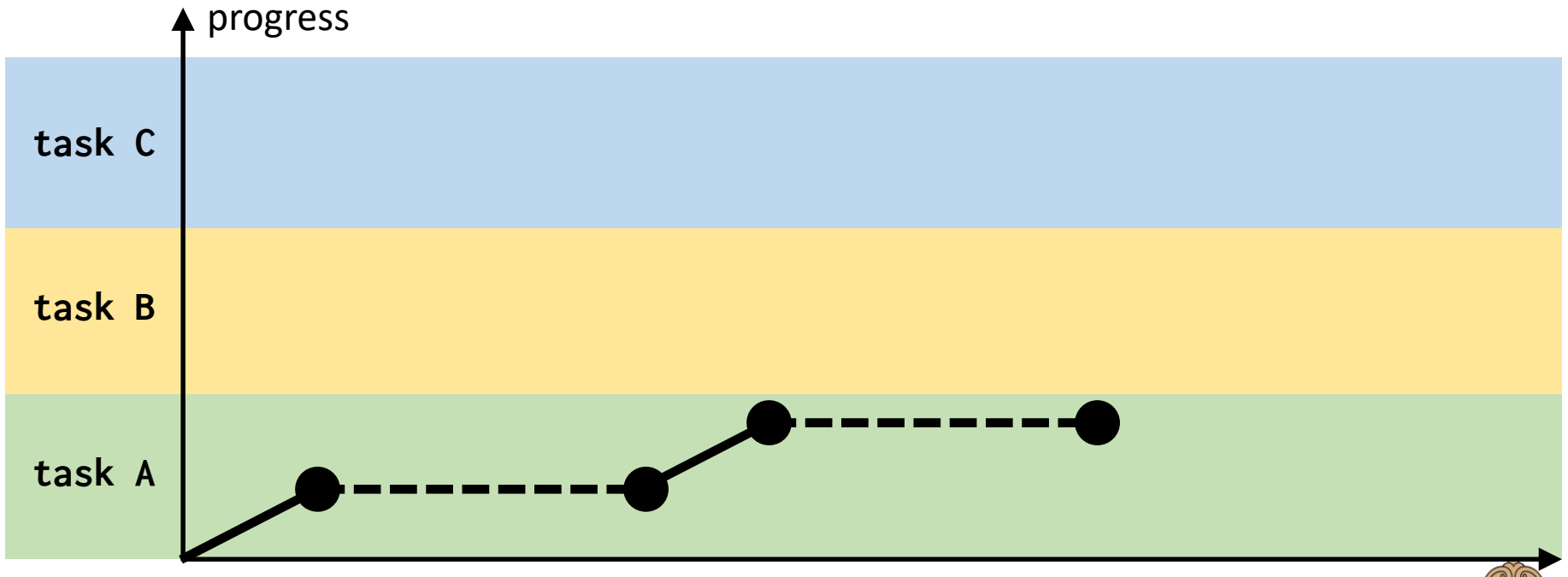
Intermittent Computing

Computation may stop at any point in the program and cannot resume until the device has harvested sufficient energy



Intermittent Computing

Computation may stop at any point in the program and cannot resume until the device has harvested sufficient energy

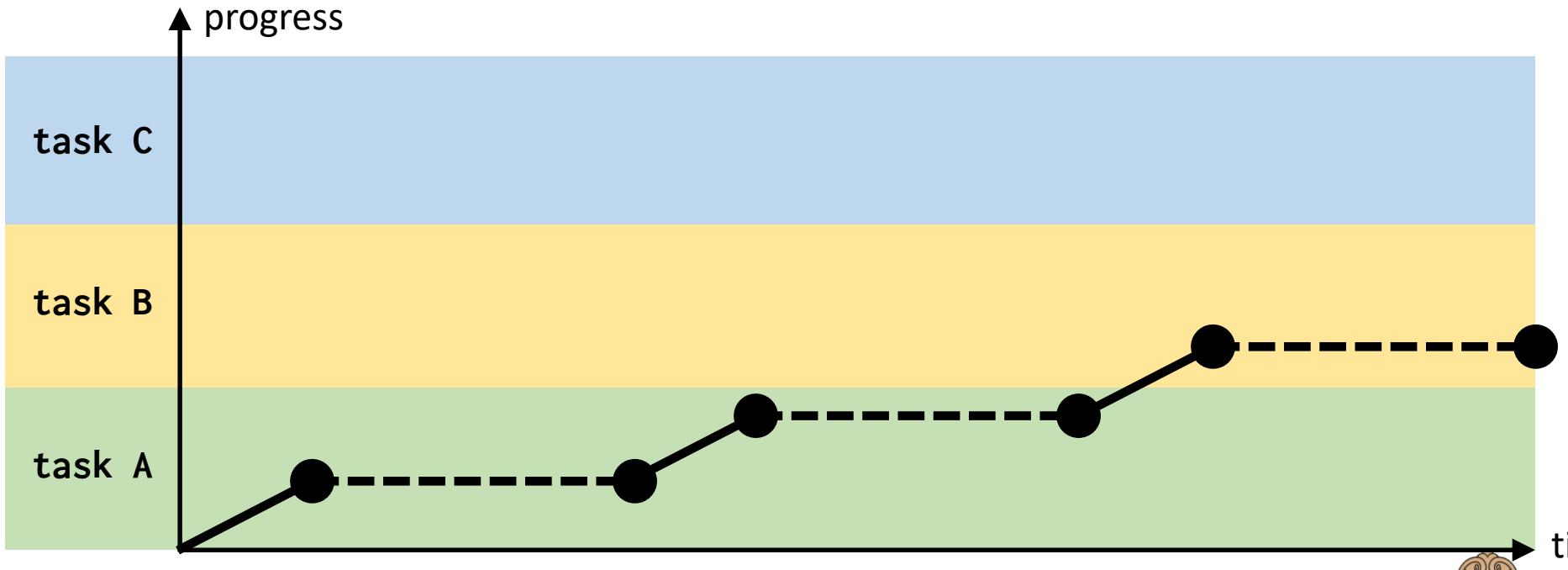


Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

Intermittent Computing

Computation may stop at any point in the program and cannot resume until the device has harvested sufficient energy

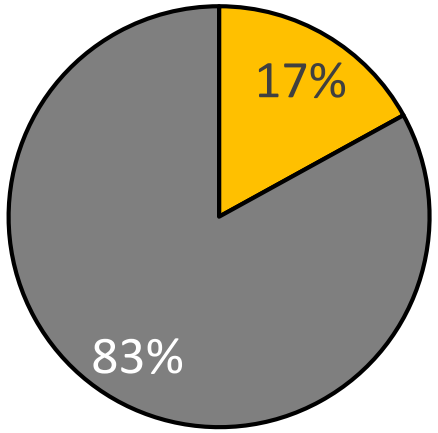
- **Power losses and backup overheads greatly impede forward progress**



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

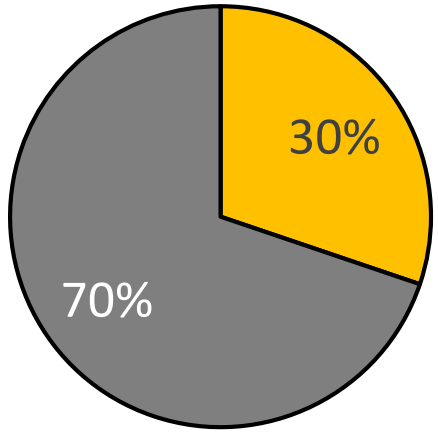
The Necessary Burden of Backups

Clank [ISCA'17]



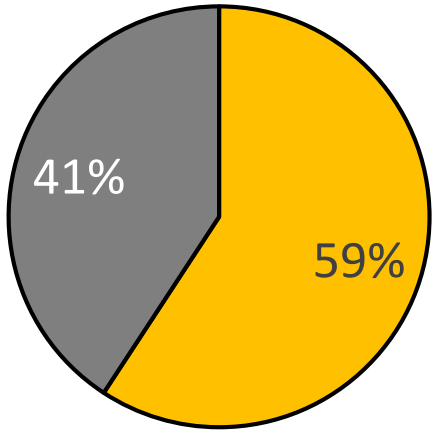
compute backup-restore

DINO [PLDI'15]



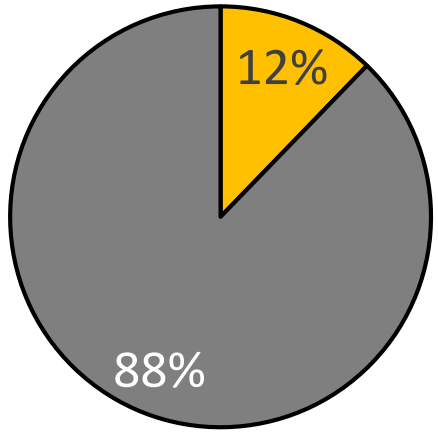
compute backup-restore

Mementos [ASPLOS'11]



compute backup-restore

NVP [HPCA'15]



compute backup-restore

Architecting for Intermittence?

Design Tools:

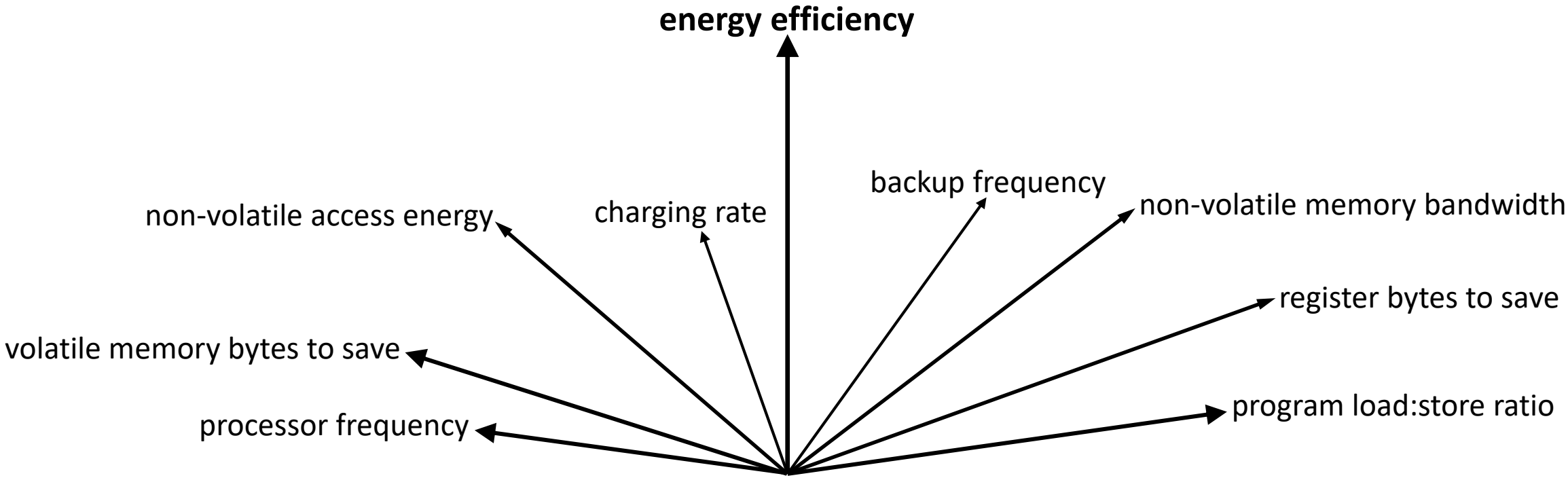
- EH Model [MICRO'18]

Design Paradigms:

- Computational Skimming [HPCA'19]

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Design Space of Intermittent Systems



And many more design axes…

- e.g., SW vs. HW backups, volatile vs. non-volatile registers, dirty vs. non-dirty data

The EH Model

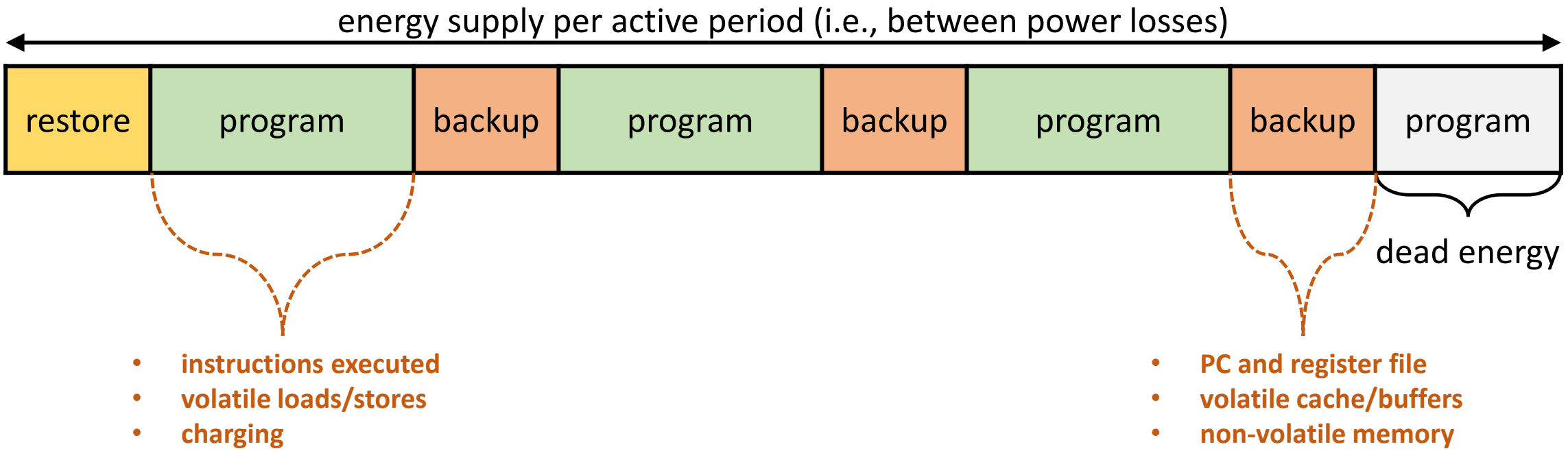
Analytical model for rapid design space exploration of arbitrary intermittent IoT systems [MICRO'18]:

- Estimates **forward progress**: % harvested energy spent on useful work
- Models significant factors that affect energy consumption:
  - Dead cycles: instructions executed that are not saved prior to power loss
  - Charging rate and capacitor size
  - Architectural state (e.g., register file) and application state (e.g., volatile data) per backup
  - Non-volatile memory latency and energy per access
  - Backup frequency; multi-backup vs. single-backup systems

The EH Model

Multi-backup systems:

- e.g., Clank [ISCA'17], Alpaca [OOPSLA'17], Mementos [ASPLOS'11]

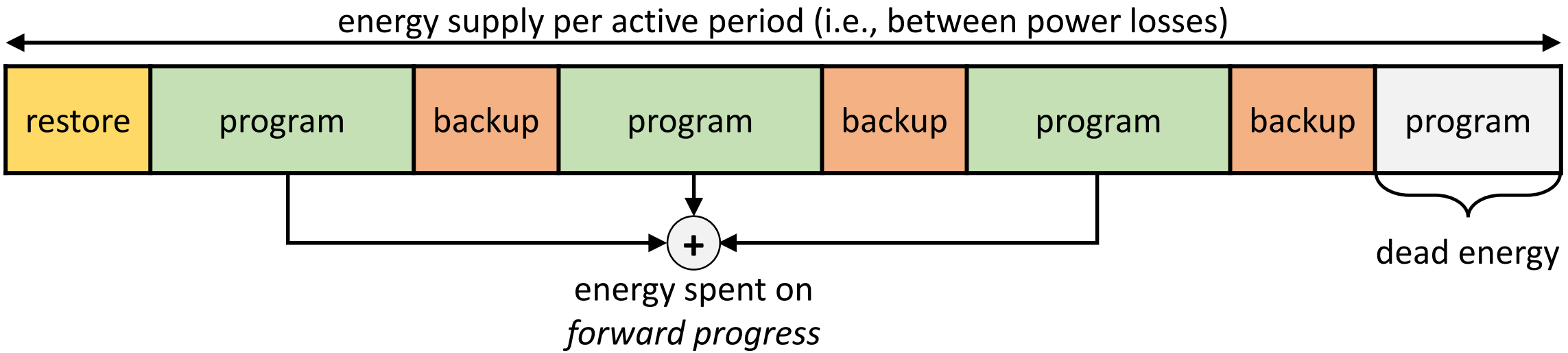


Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

The EH Model

Multi-backup systems:

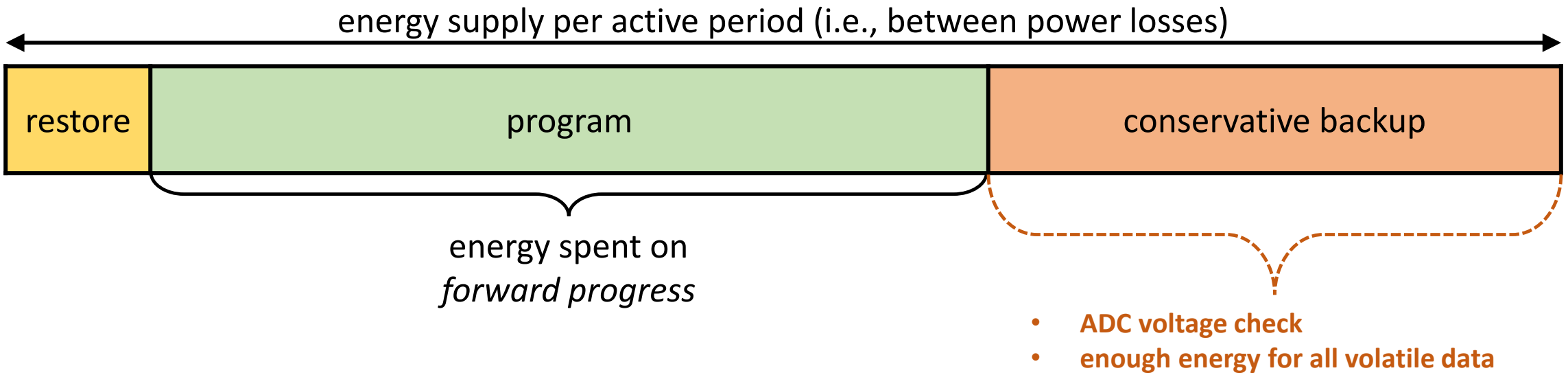
- e.g., Clank [ISCA'17], Alpaca [OOPSLA'17], Mementos [ASPLOS'11]



The EH Model

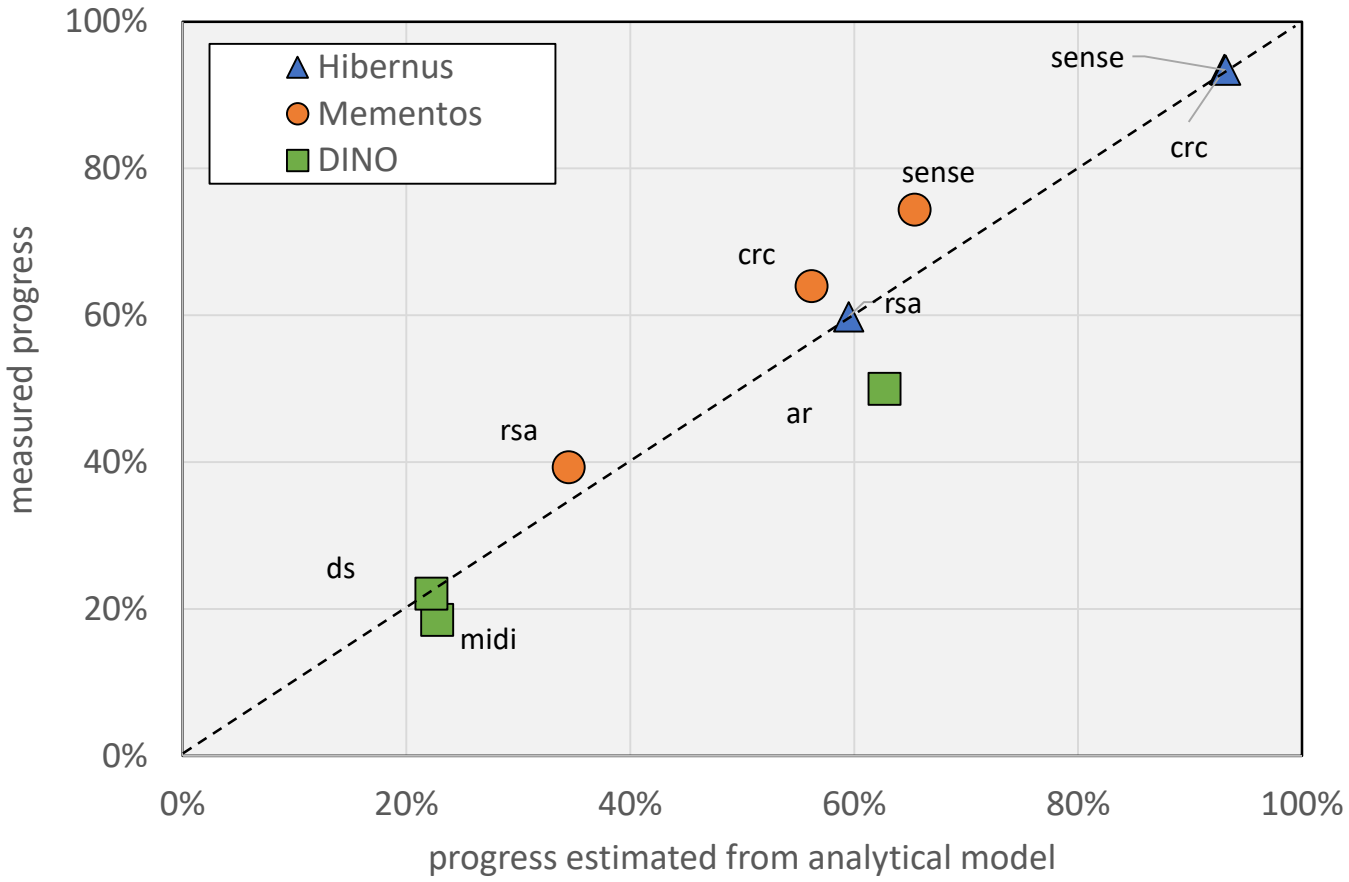
Single-backup systems:

- e.g., Hibernus [IEEE TCAD'16], NVP [HPCA'15], QuickRecall [VLSID'14]



The EH Model

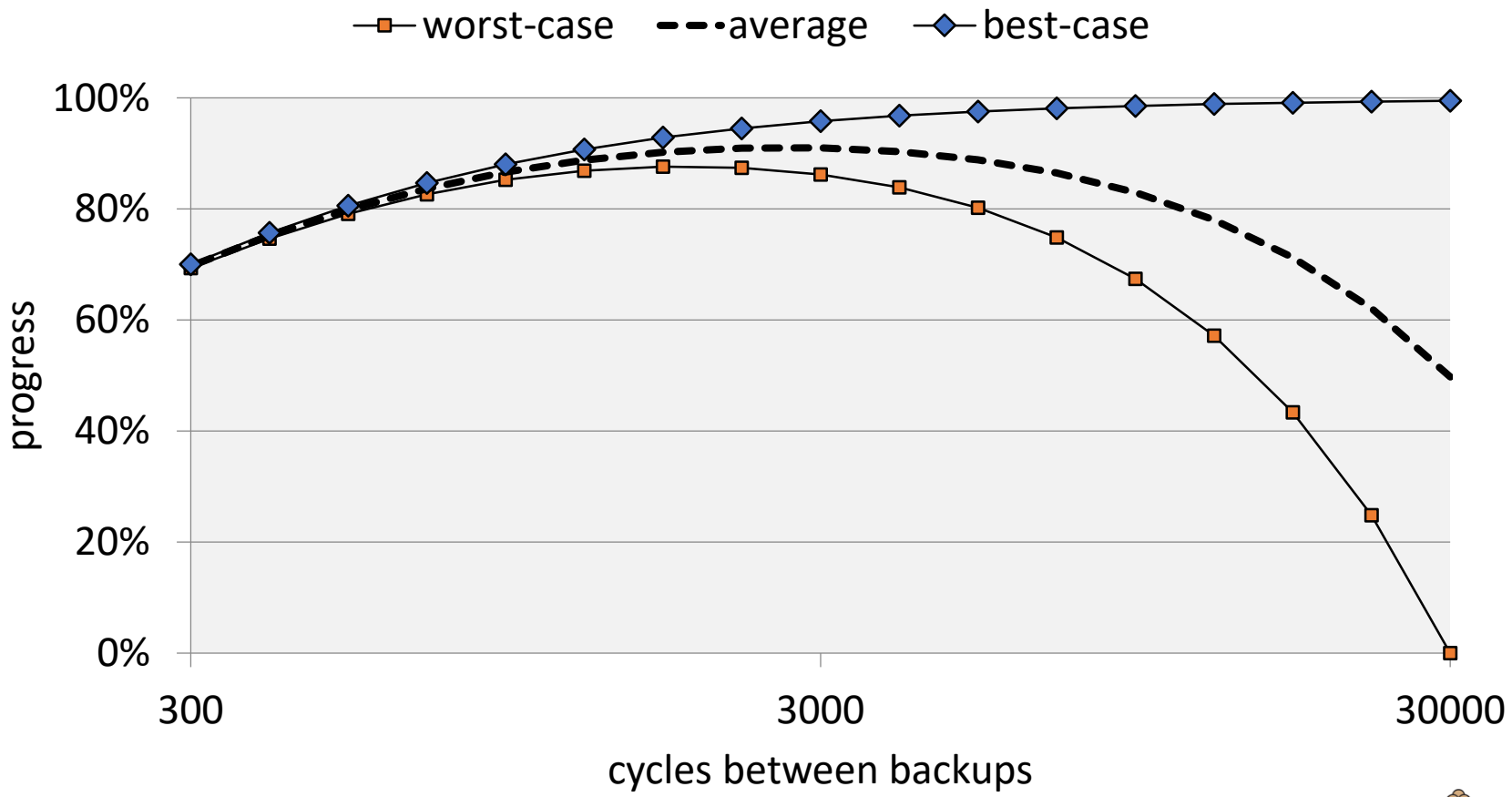
Evaluation on MSP430 LaunchPad:



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

The EH Model – Analytical Explorations

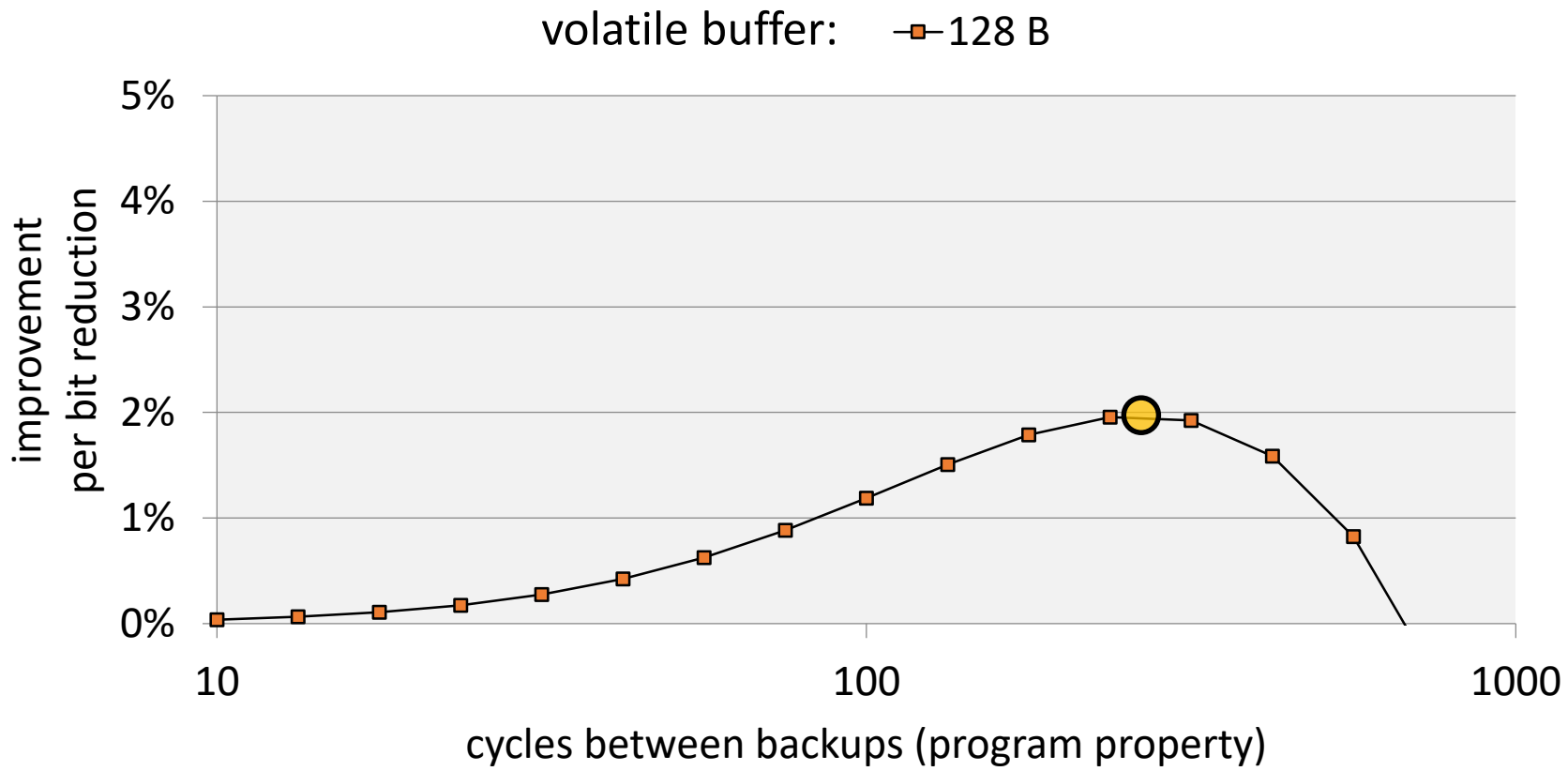
e.g., Mementos [ASPLOS'11]:



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

The EH Model – Analytical Explorations

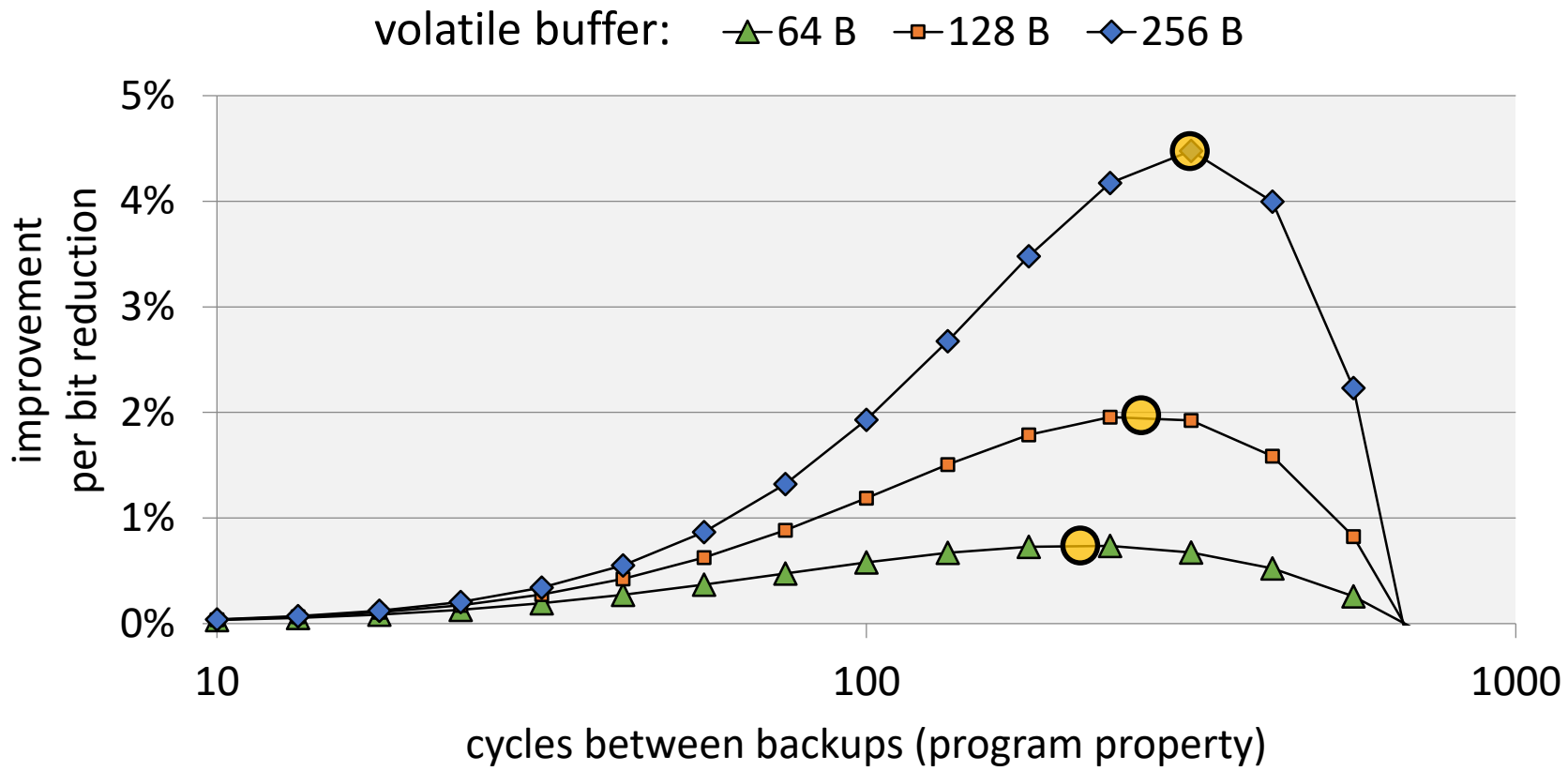
e.g., Reduced-precision image recognition on Clank [ISCA'17]:



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

The EH Model – Analytical Explorations

e.g., Reduced-precision image recognition on Clank [ISCA'17]:



The EH Model – Analytical Explorations

More case studies and explorations [MICRO'18]:

- Saving architectural vs. application state
- The benefits of store-major locality
- Controlling backups via write-after-read dependences

Architecting for Intermittence?

Design Tools:

- EH Model [MICRO'18]

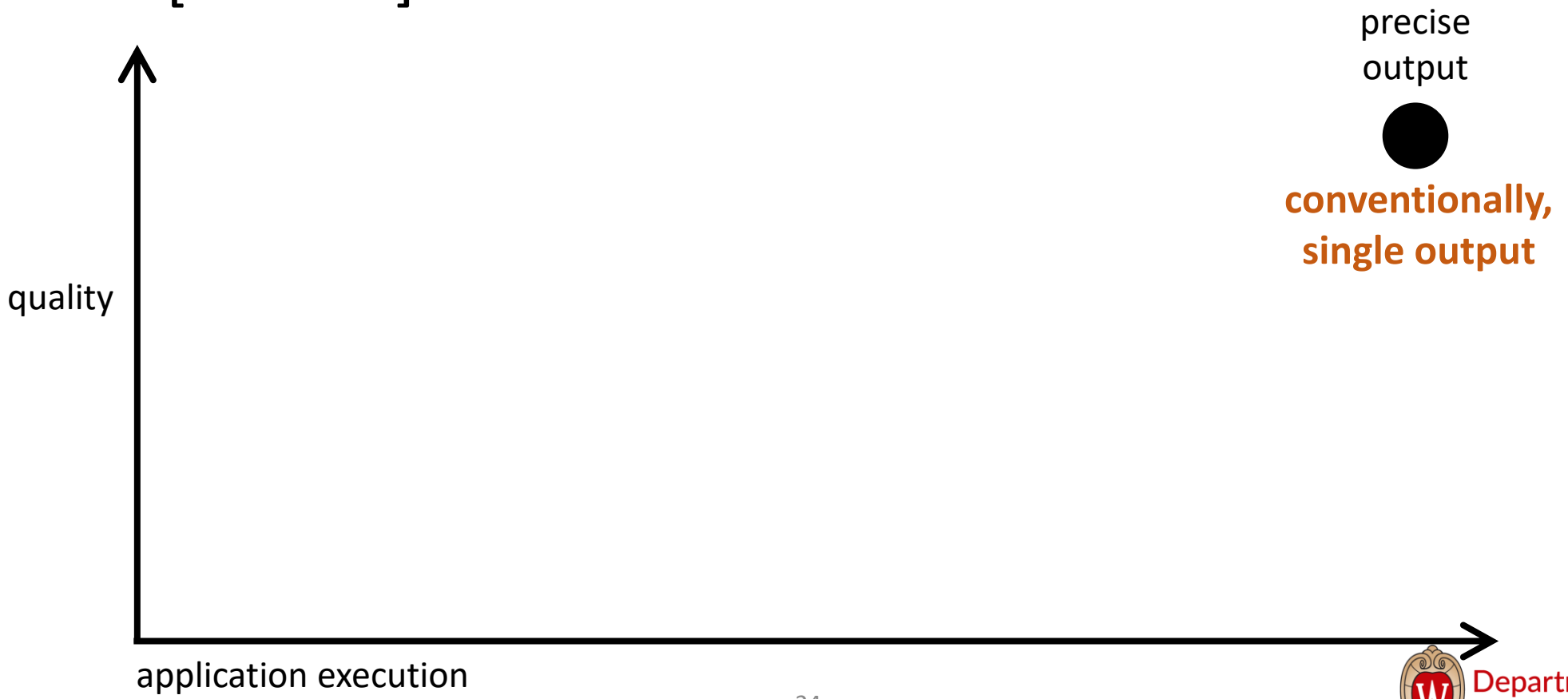
Design Paradigms:

- **Computational Skimming [HPCA'19]**

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

The Anytime Automaton Model

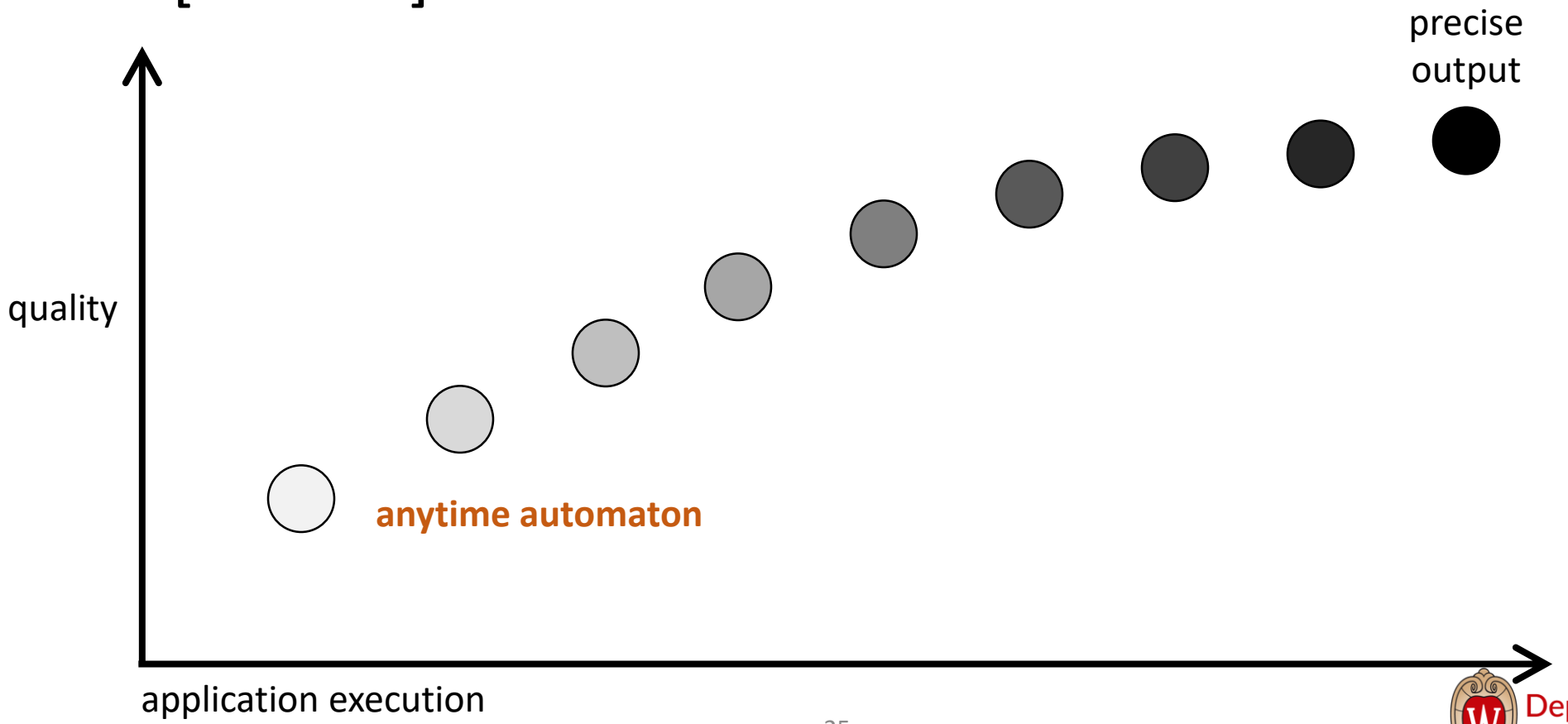
General-purpose computation model for quality-proportional execution [ISCA'16]:



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

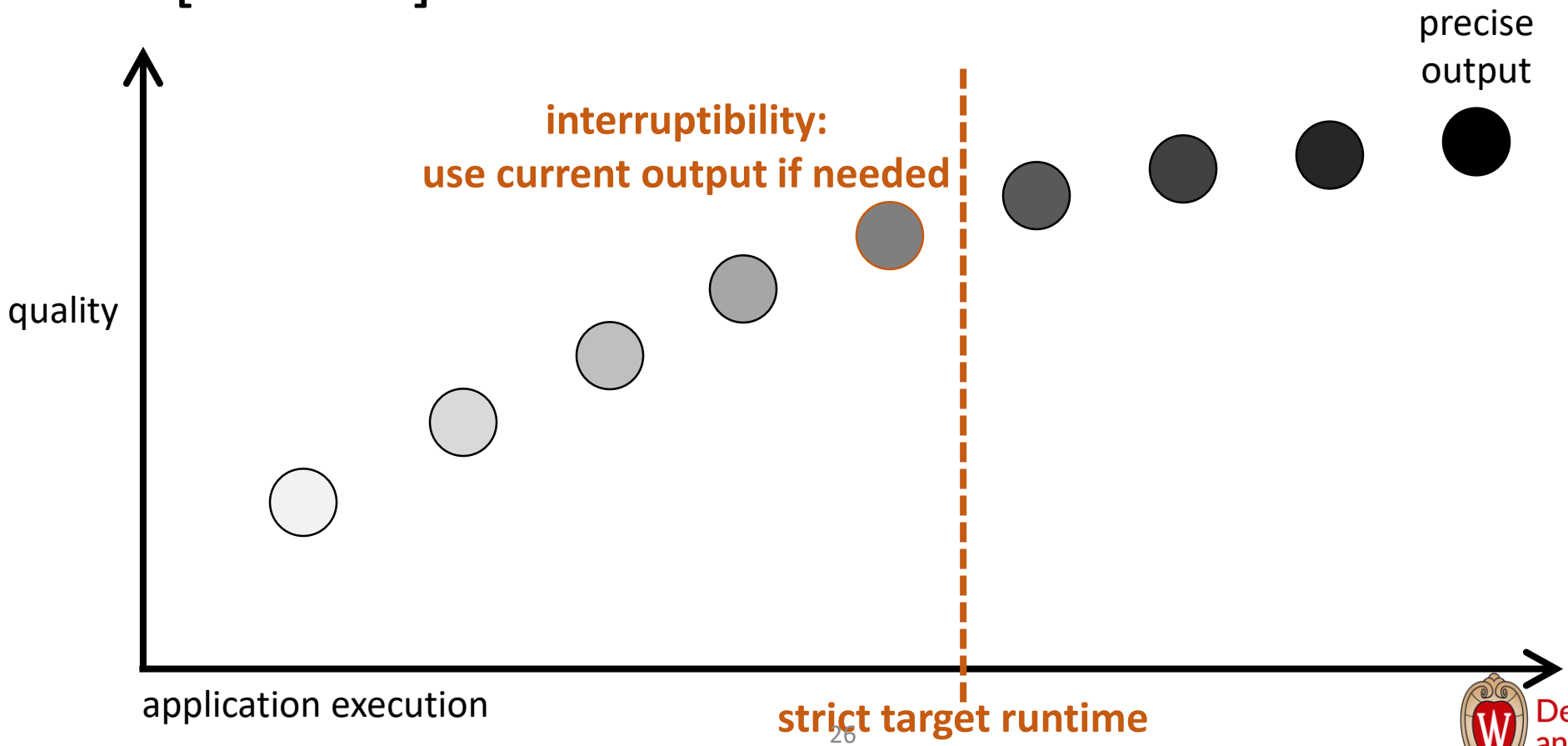
The Anytime Automaton Model

General-purpose computation model for quality-proportional execution [ISCA'16]:



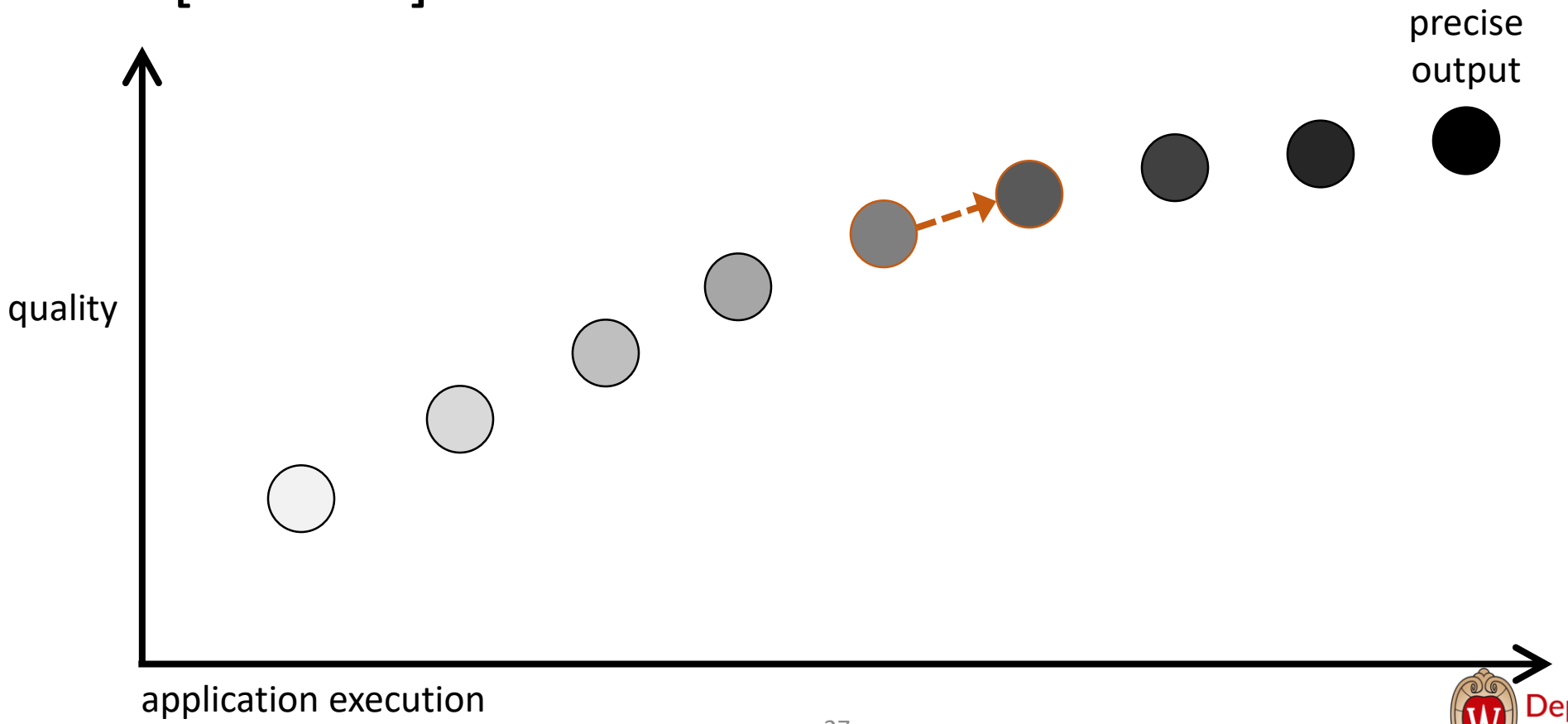
The Anytime Automaton Model

General-purpose computation model for quality-proportional execution [ISCA'16]:



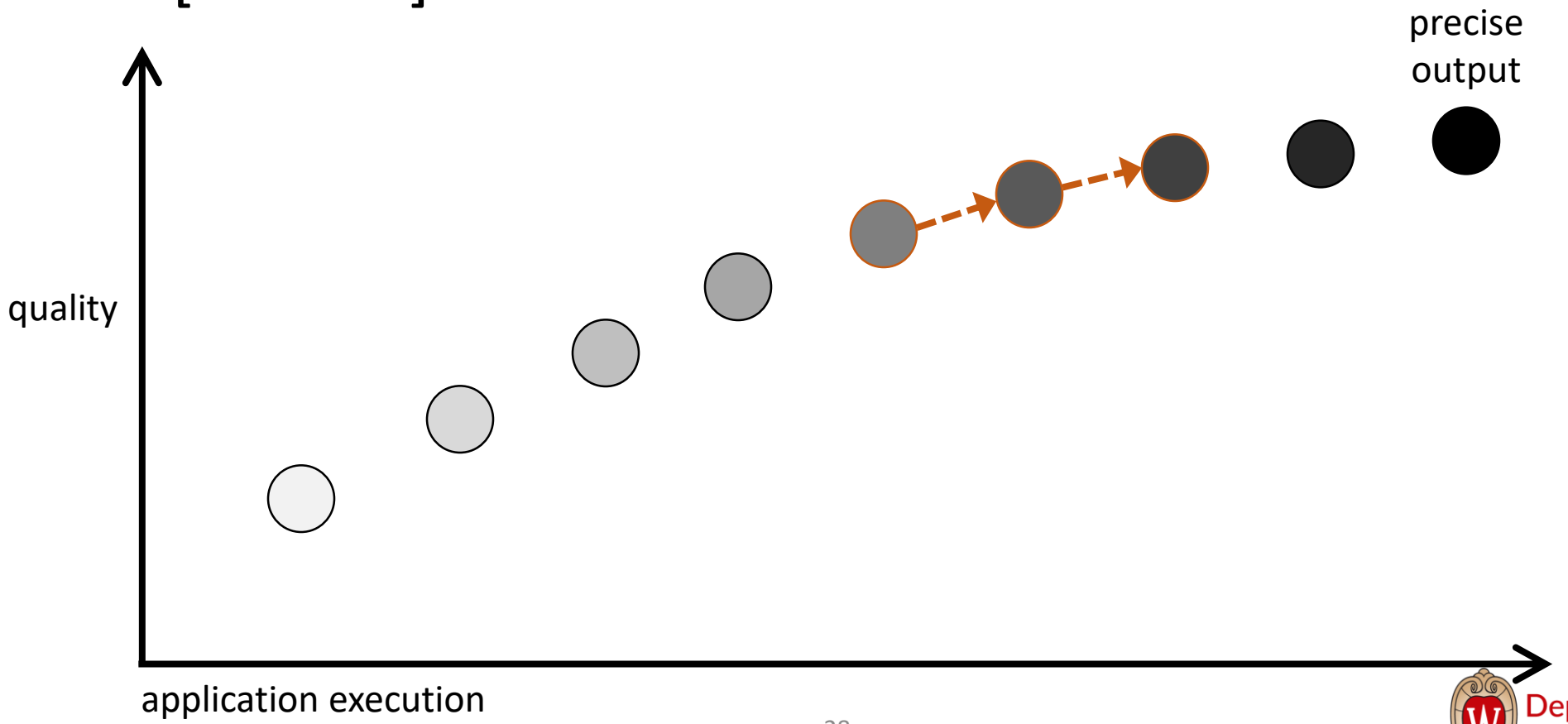
The Anytime Automaton Model

General-purpose computation model for quality-proportional execution [ISCA'16]:



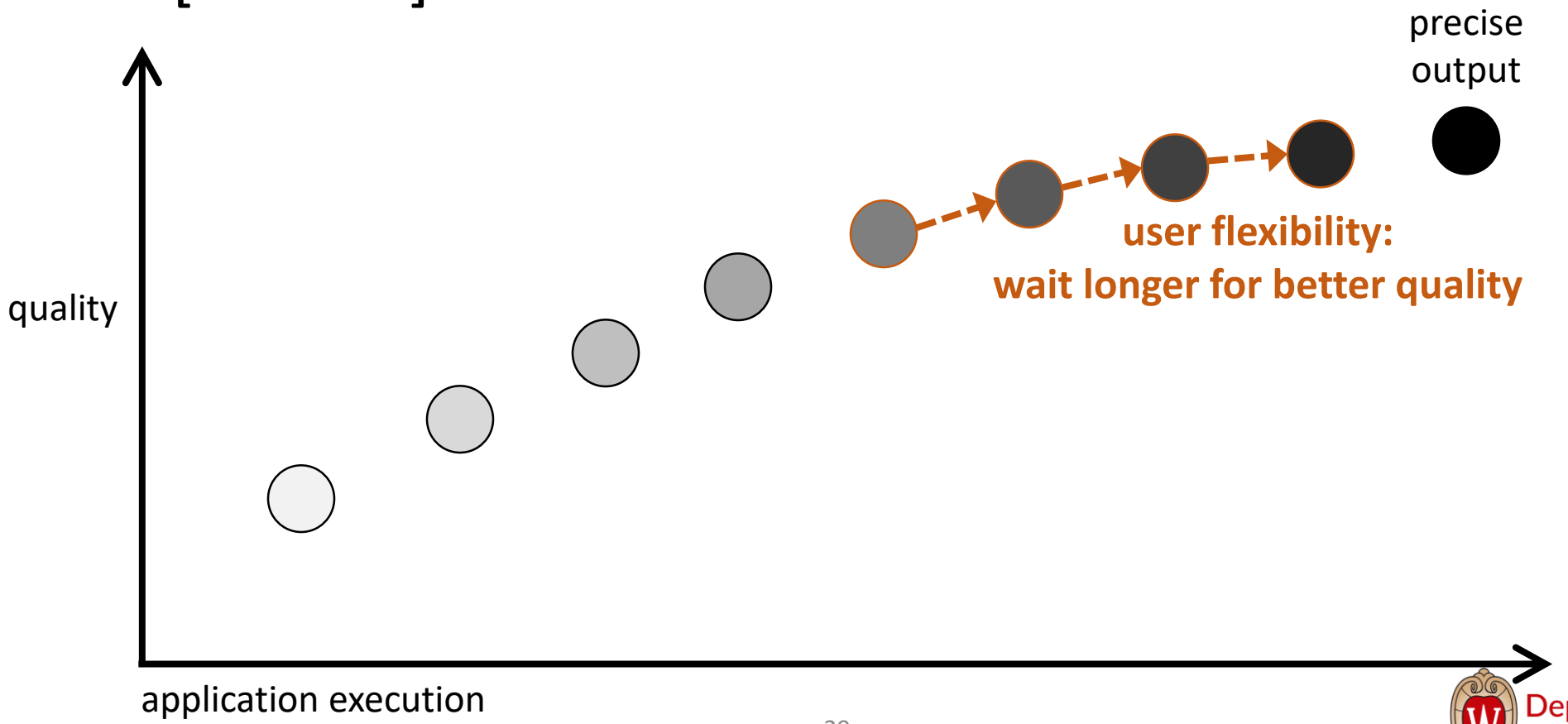
The Anytime Automaton Model

General-purpose computation model for quality-proportional execution [ISCA'16]:



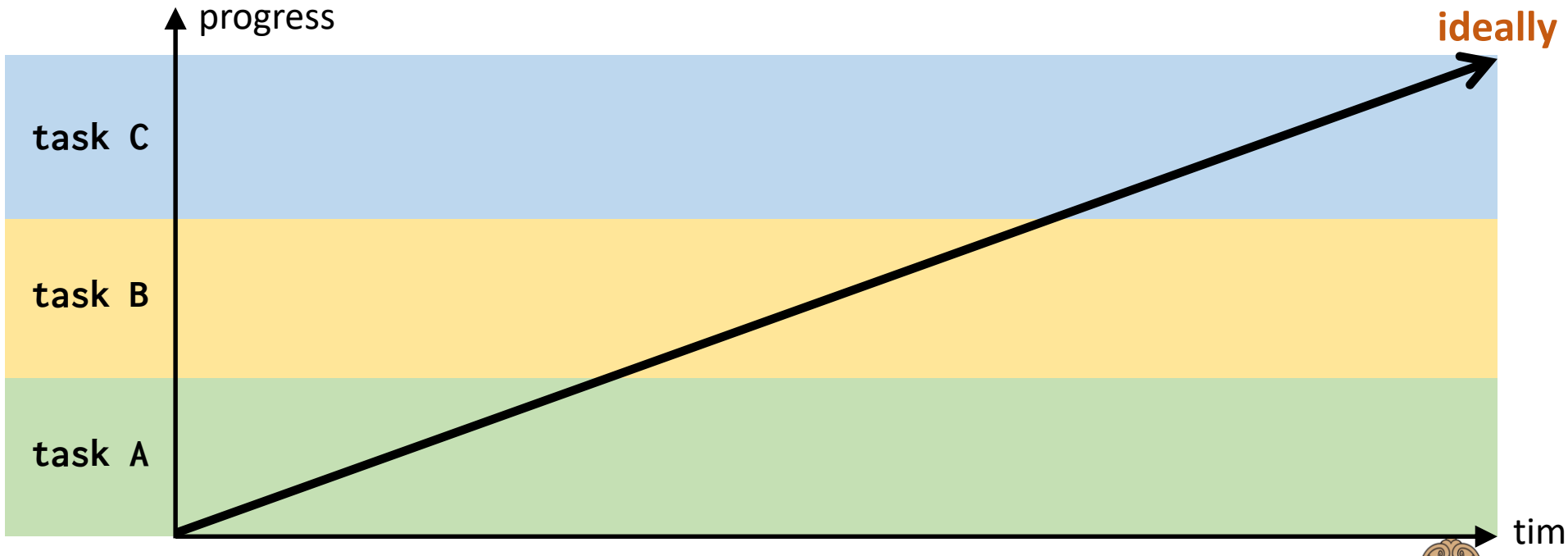
The Anytime Automaton Model

General-purpose computation model for quality-proportional execution [ISCA'16]:



Intermittent Computing

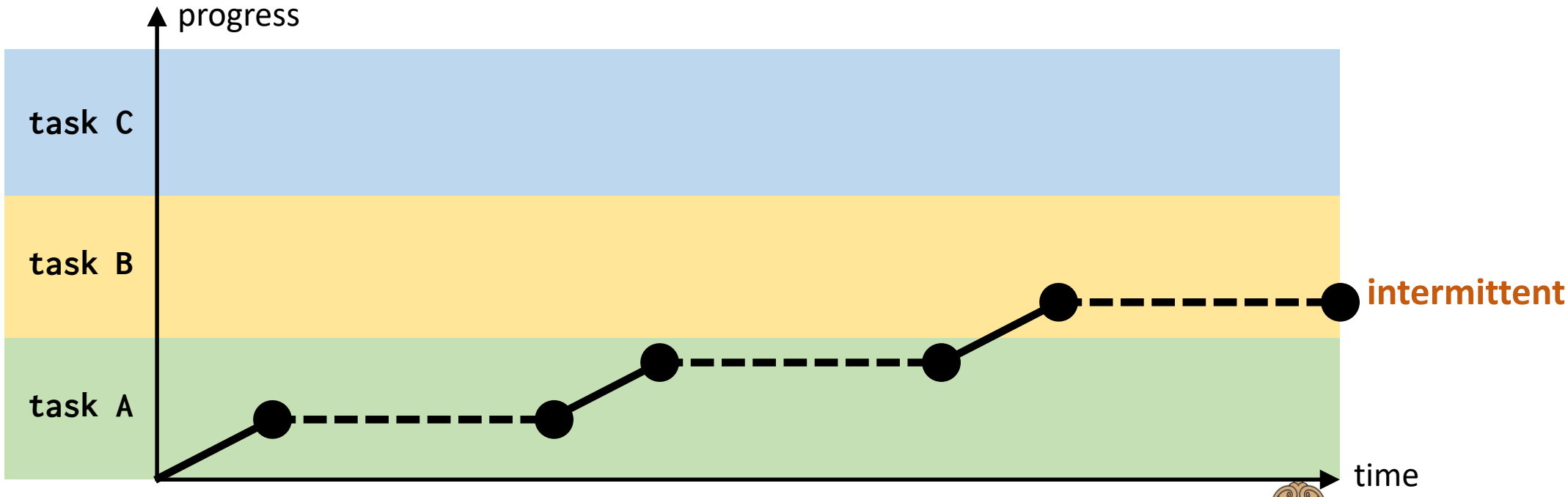
Computation may stop at any point in the program and cannot resume until the device has harvested sufficient energy



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

Intermittent Computing

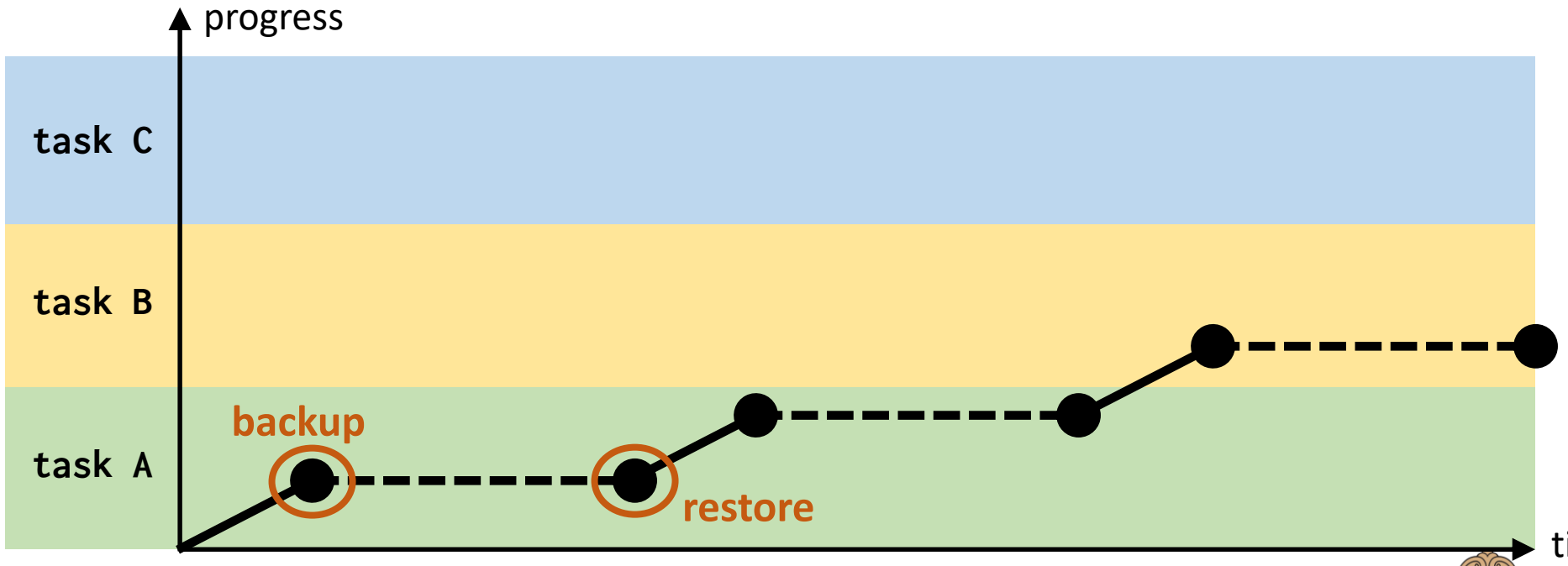
Computation may stop at any point in the program and cannot resume until the device has harvested sufficient energy



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

Computational Skimming

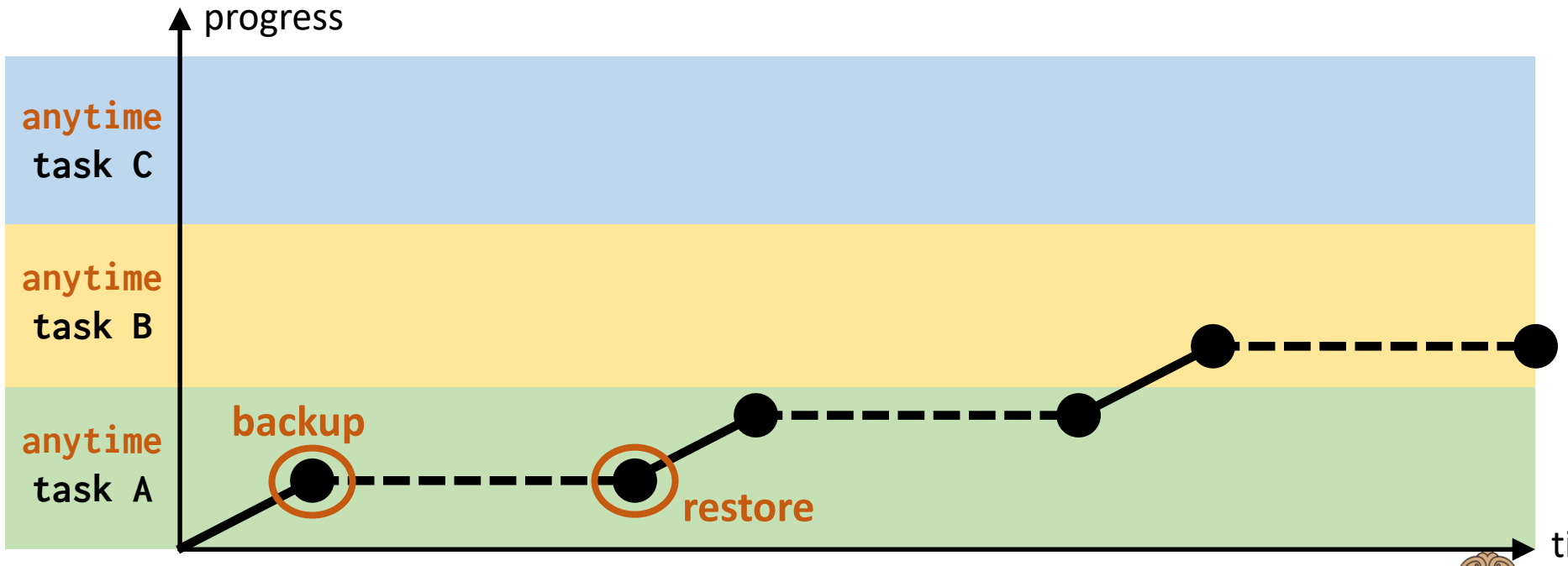
Insight: Decouple backup from restore point via anytime model



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming

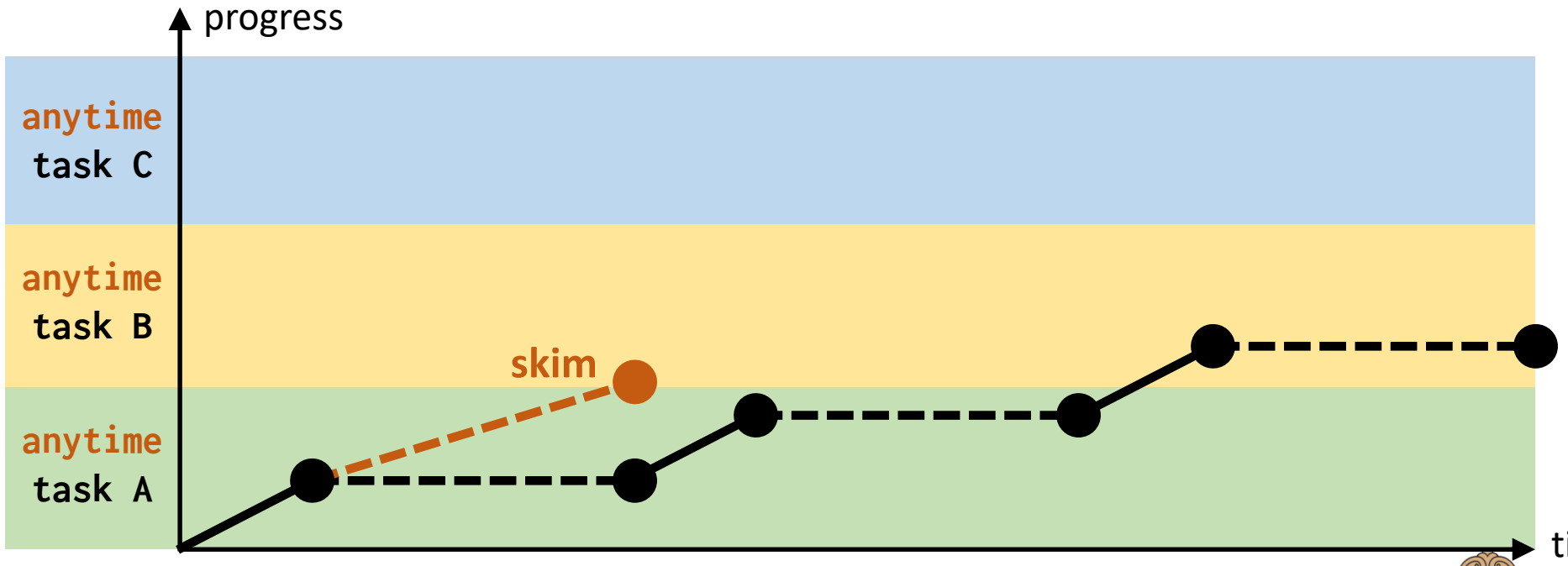
Insight: Decouple backup from restore point via anytime model



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming

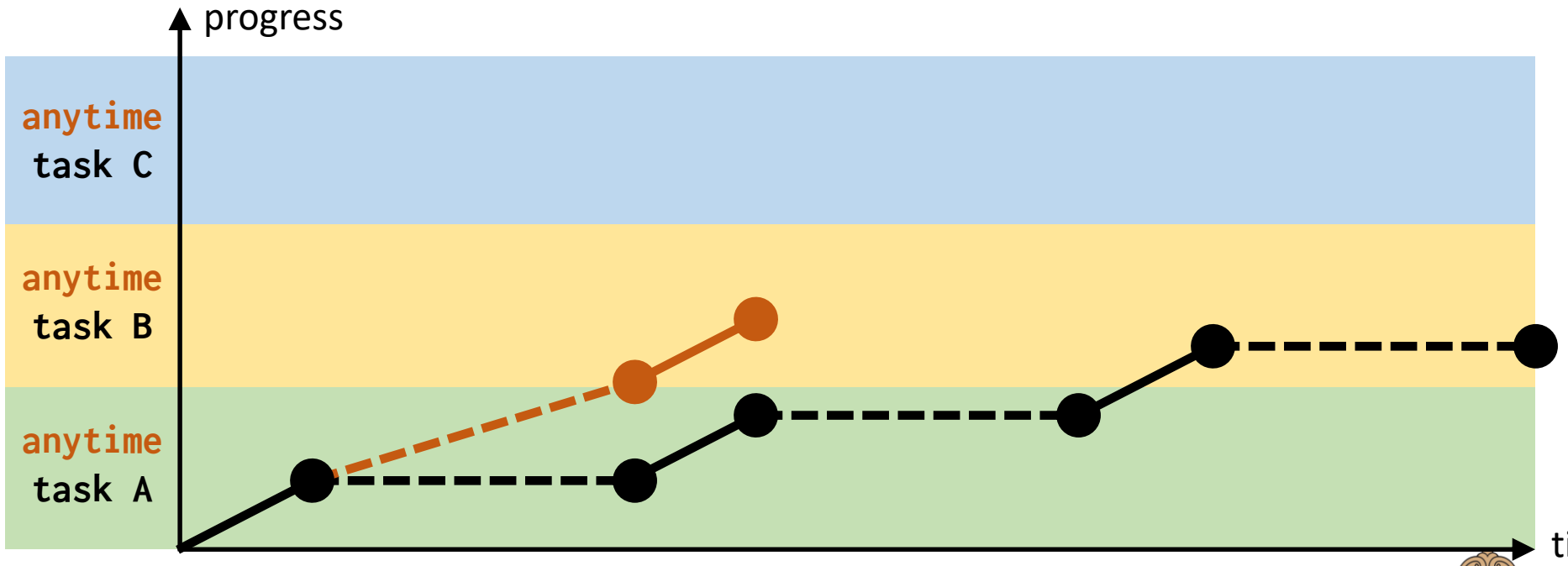
Insight: Decouple backup from restore point via anytime model



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming

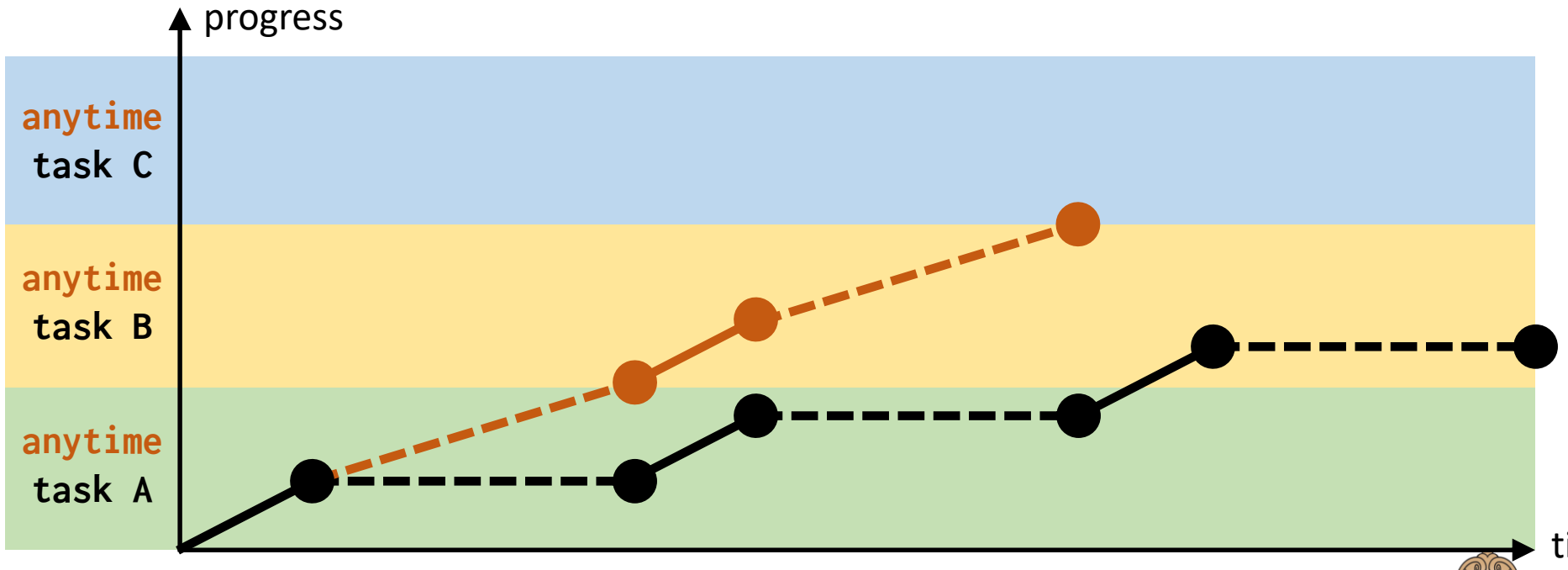
Insight: Decouple backup from restore point via anytime model



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming

Insight: Decouple backup from restore point via anytime model

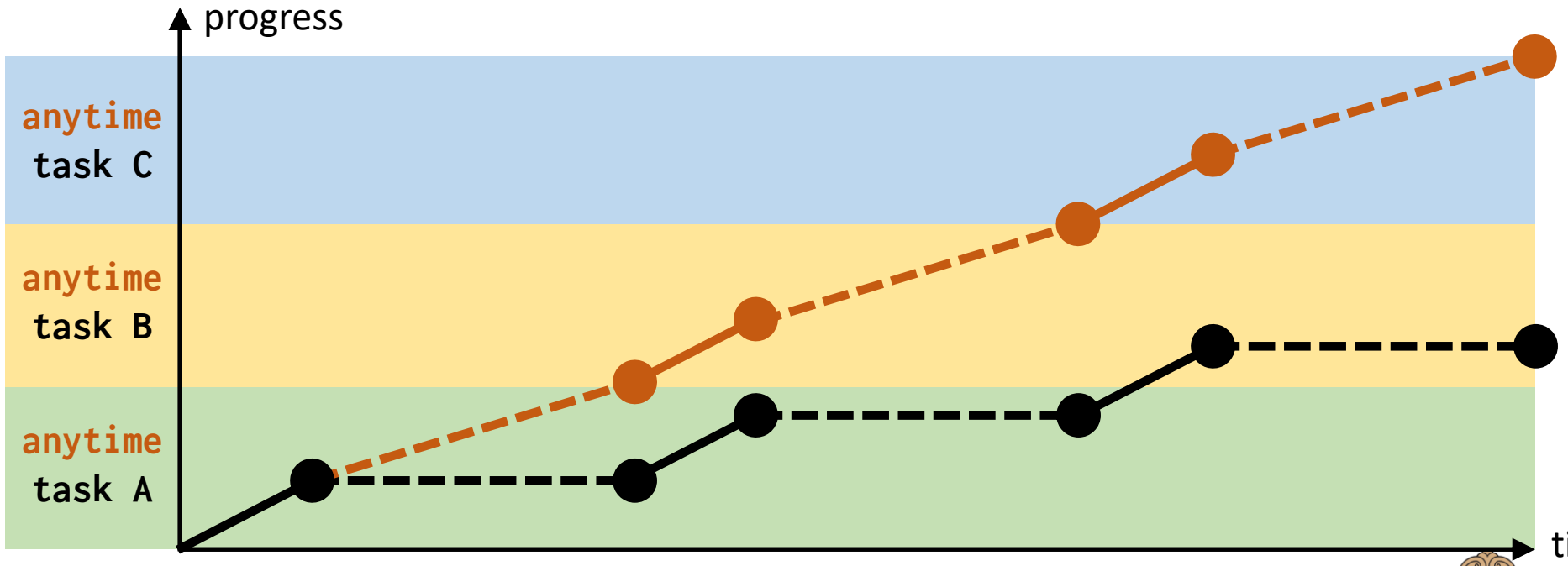


Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming

Insight: Decouple backup from restore point via anytime model

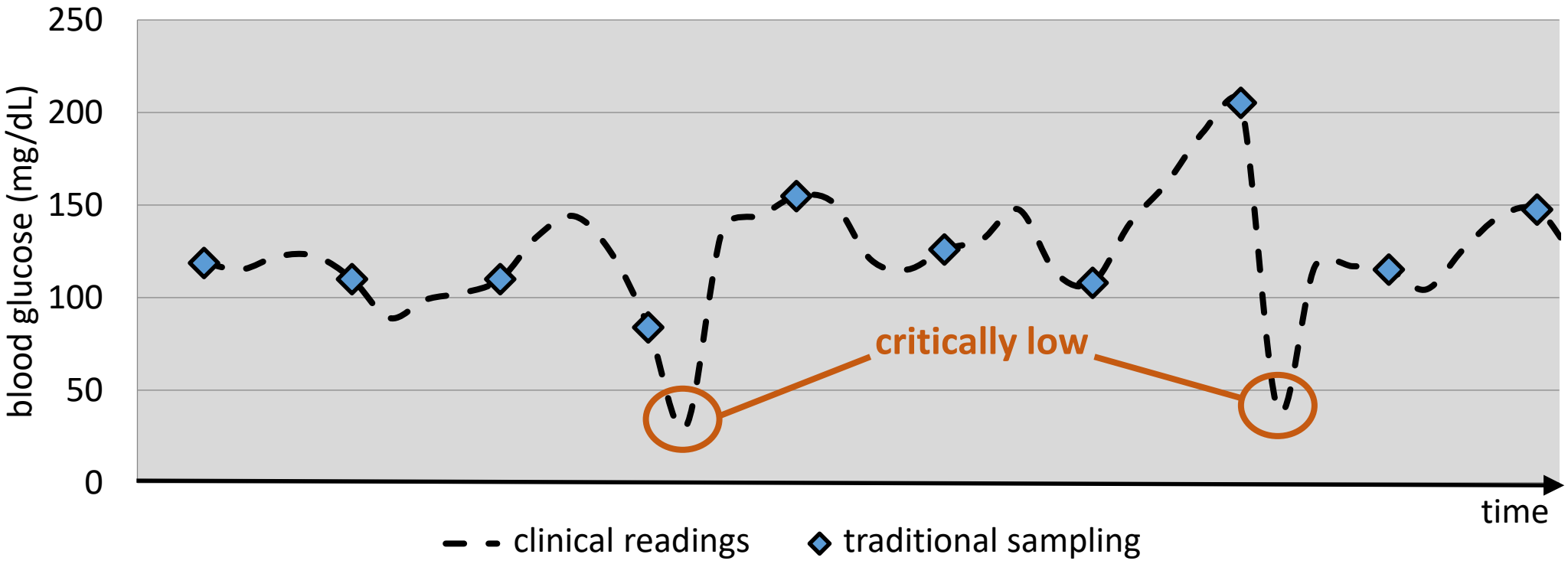
Computational Skimming: Quality of program output scales with how much energy we can afford



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

Computational Skimming

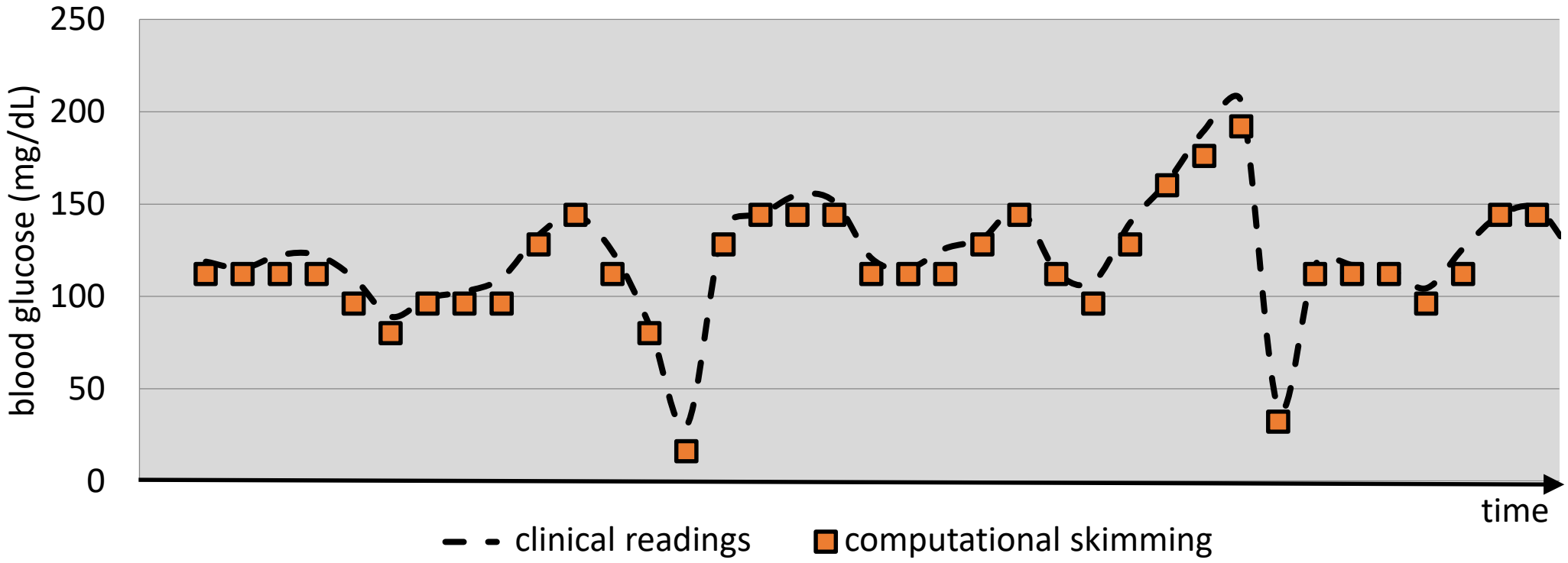
e.g., Blood glucose monitoring:



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming

e.g., Blood glucose monitoring:



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

Computational Skimming – What's Next

What's Next processor architecture for computational skimming, designed from baseline ARM M0+ [HPCA'19]:

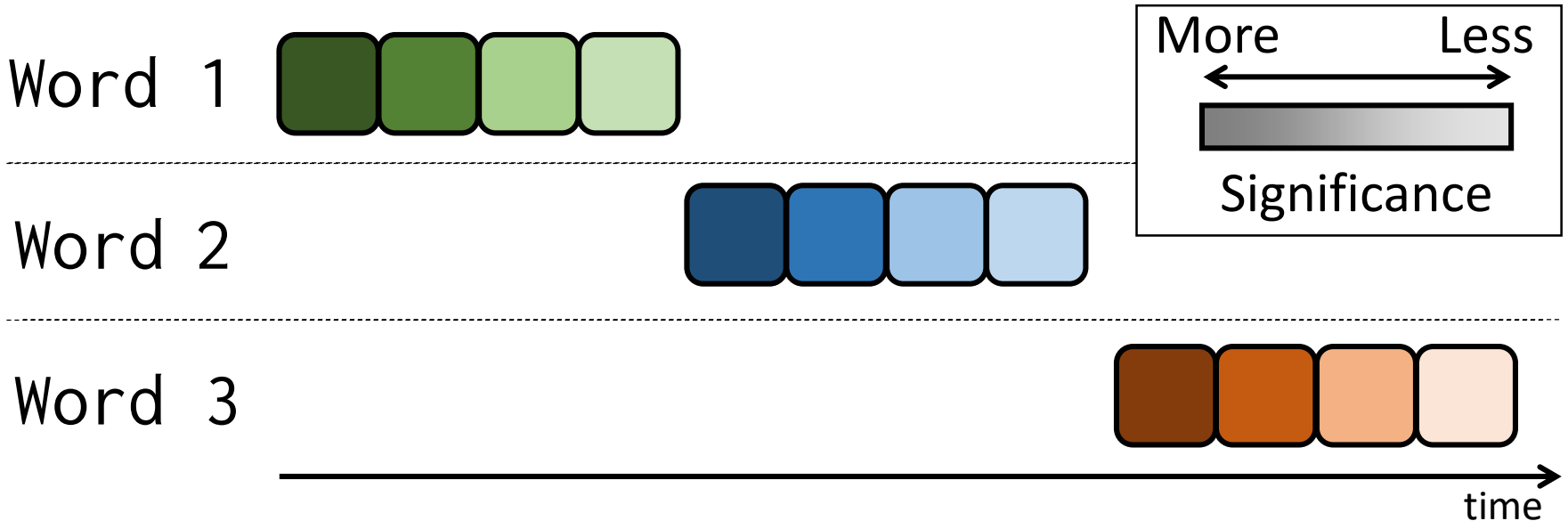
- Anytime subword pipelining
- Anytime subword vectorization
- Subword memoization

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming – Anytime Pipelining

Long-latency operations (e.g., mul):

- Conventional

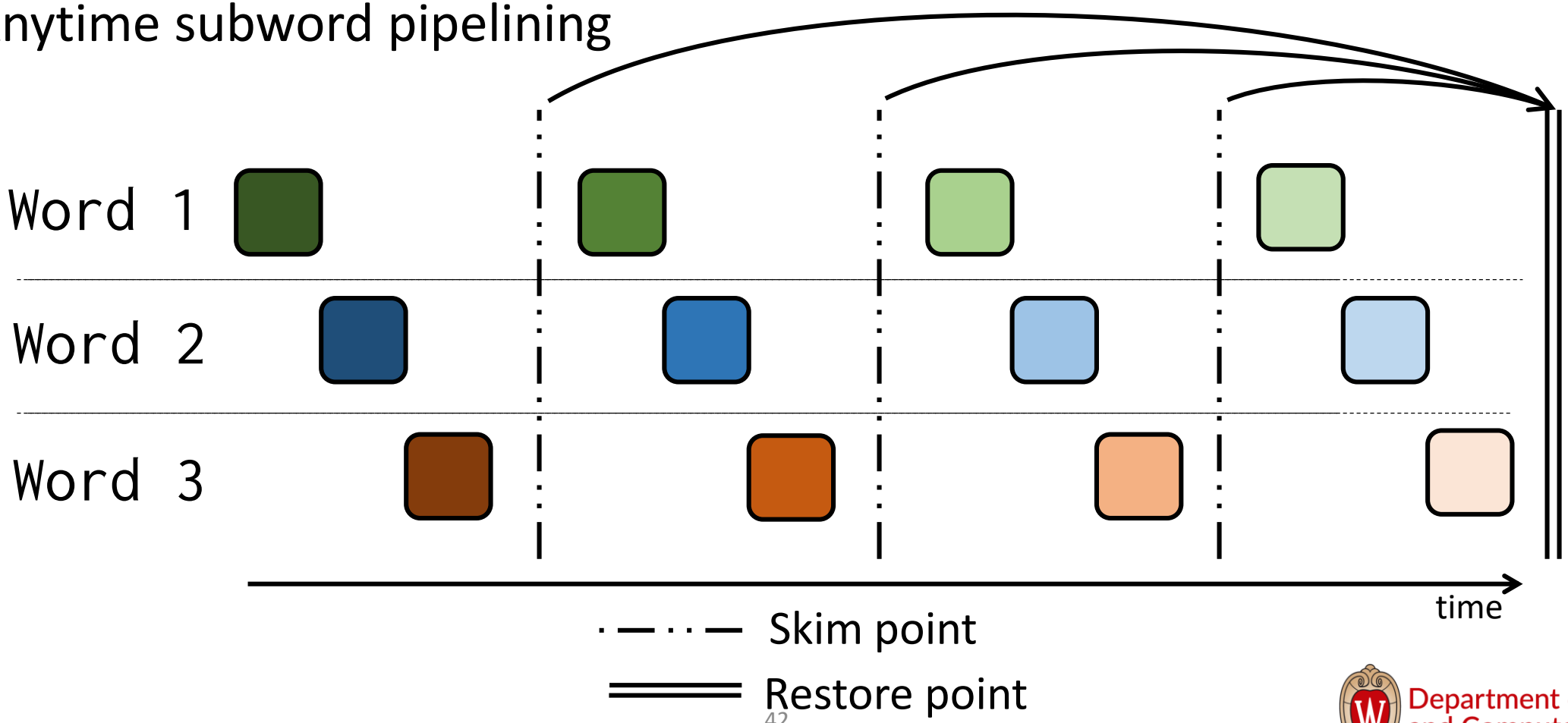


Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming – Anytime Pipelining

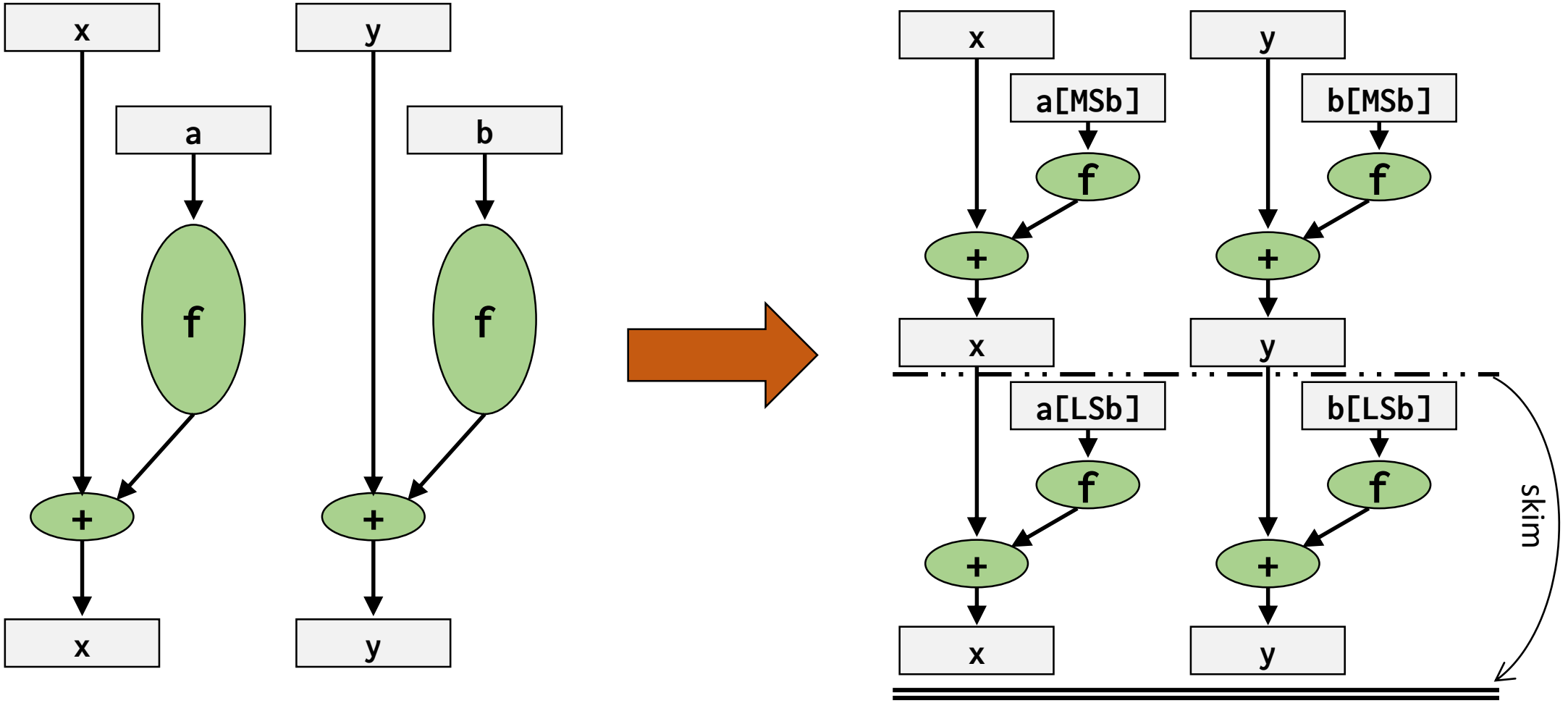
Long-latency operations (e.g., mul):

- Anytime subword pipelining



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

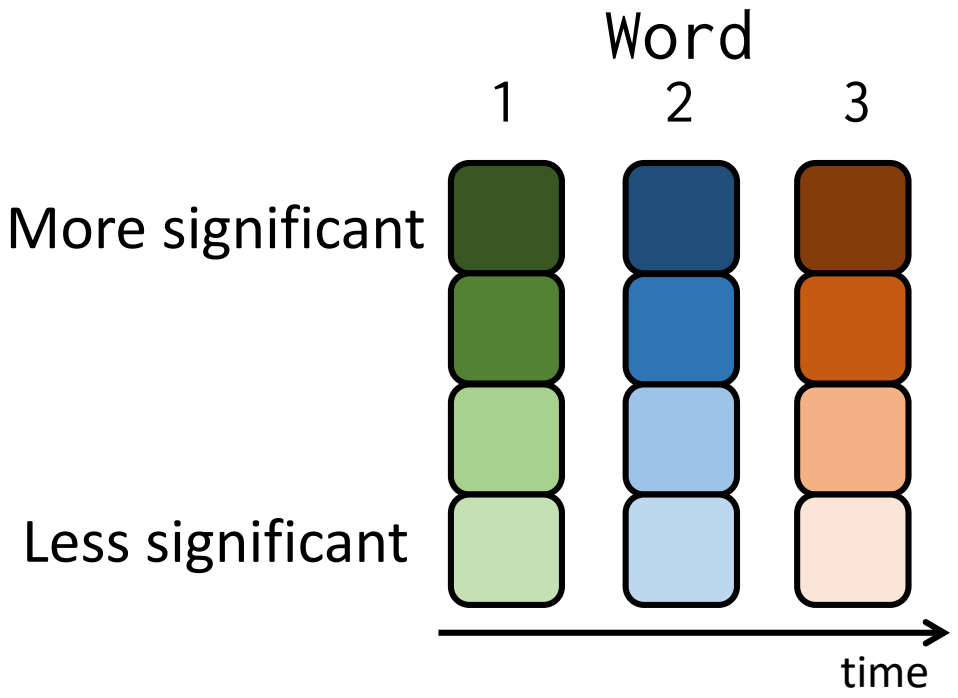
Computational Skimming – Anytime Pipelining



Computational Skimming – Anytime Vectorization

Short-latency operations (e.g., add):

- Conventional

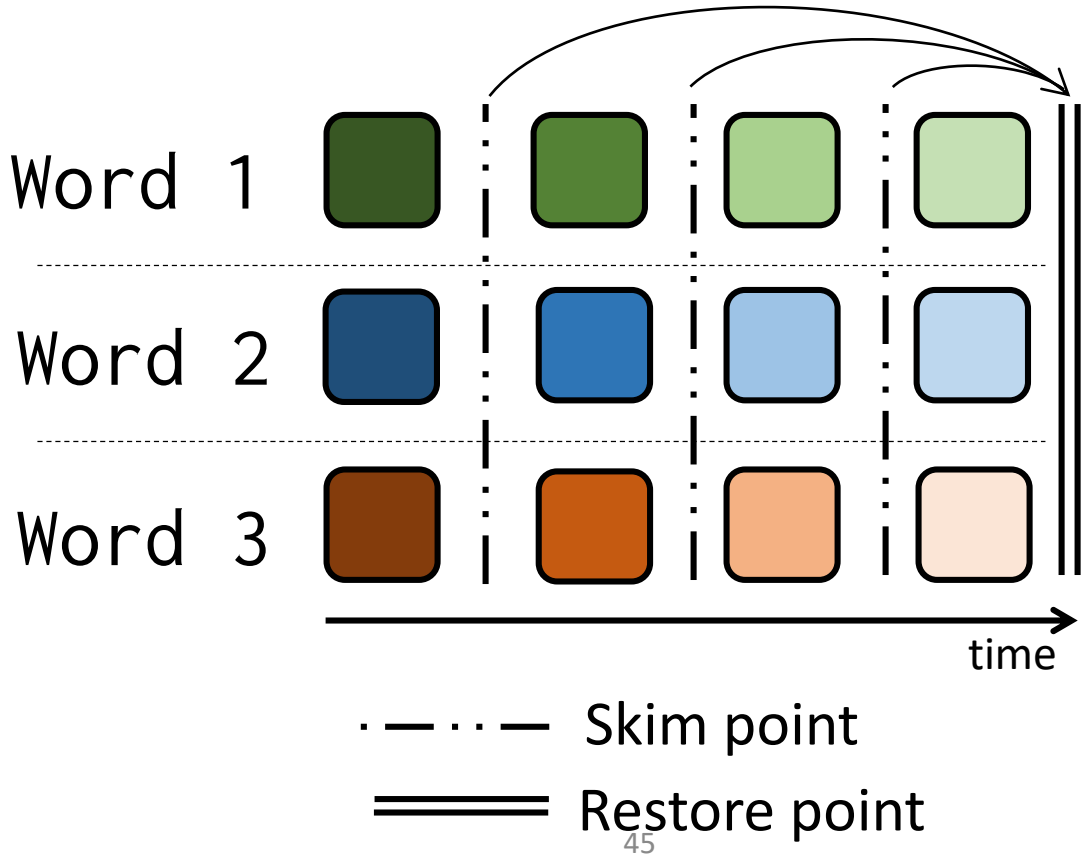


Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming – Anytime Vectorization

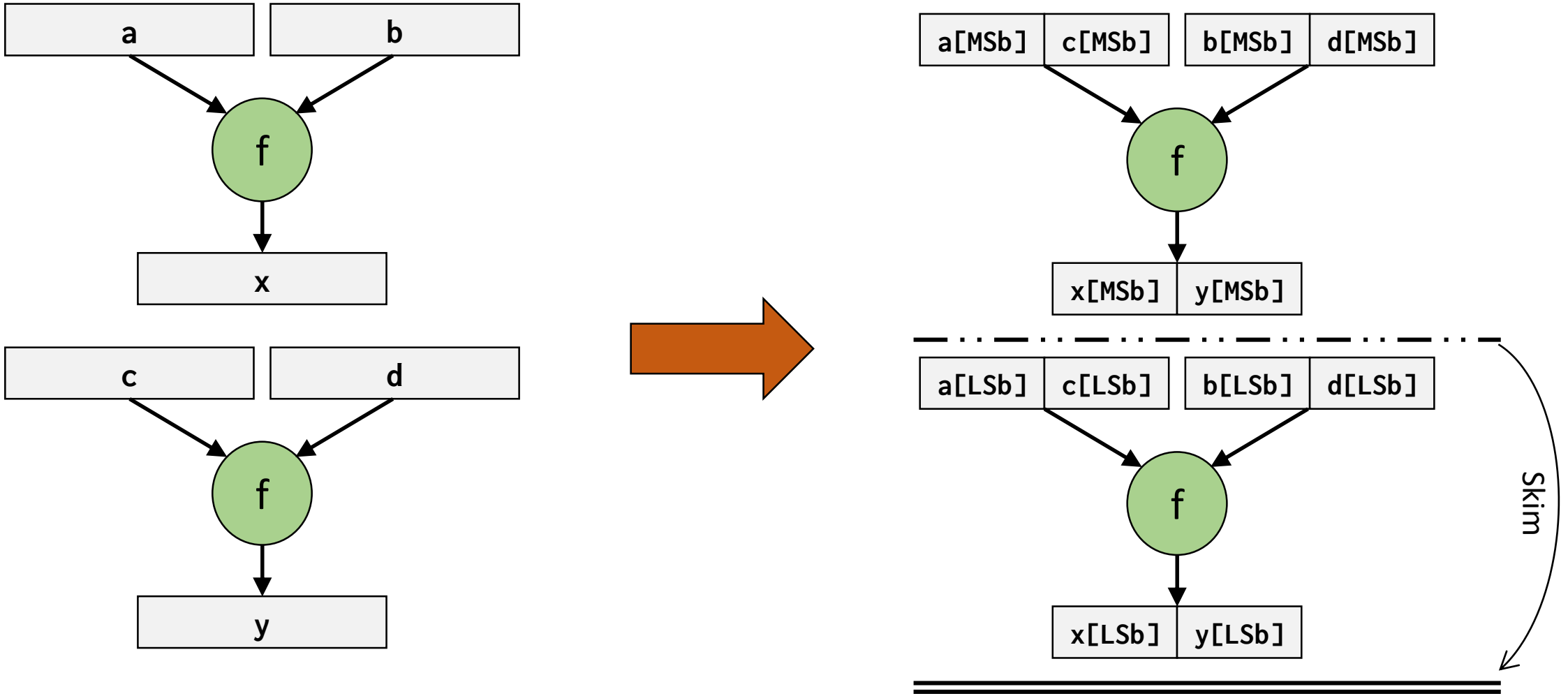
Short-latency operations (e.g., add):

- Anytime subword vectorization



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

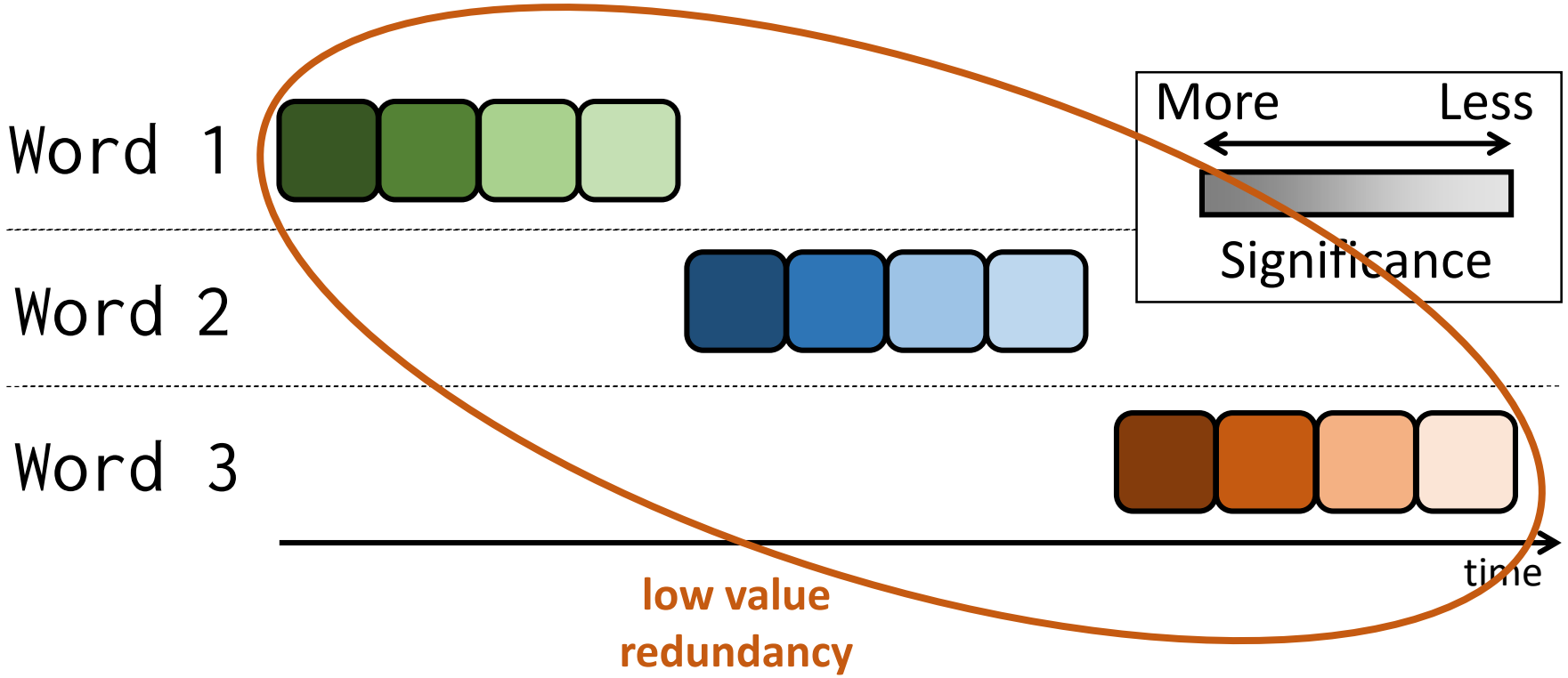
Computational Skimming – Anytime Vectorization



Computational Skimming – Subword Memoization

In conjunction with anytime subword pipelining:

- Conventional

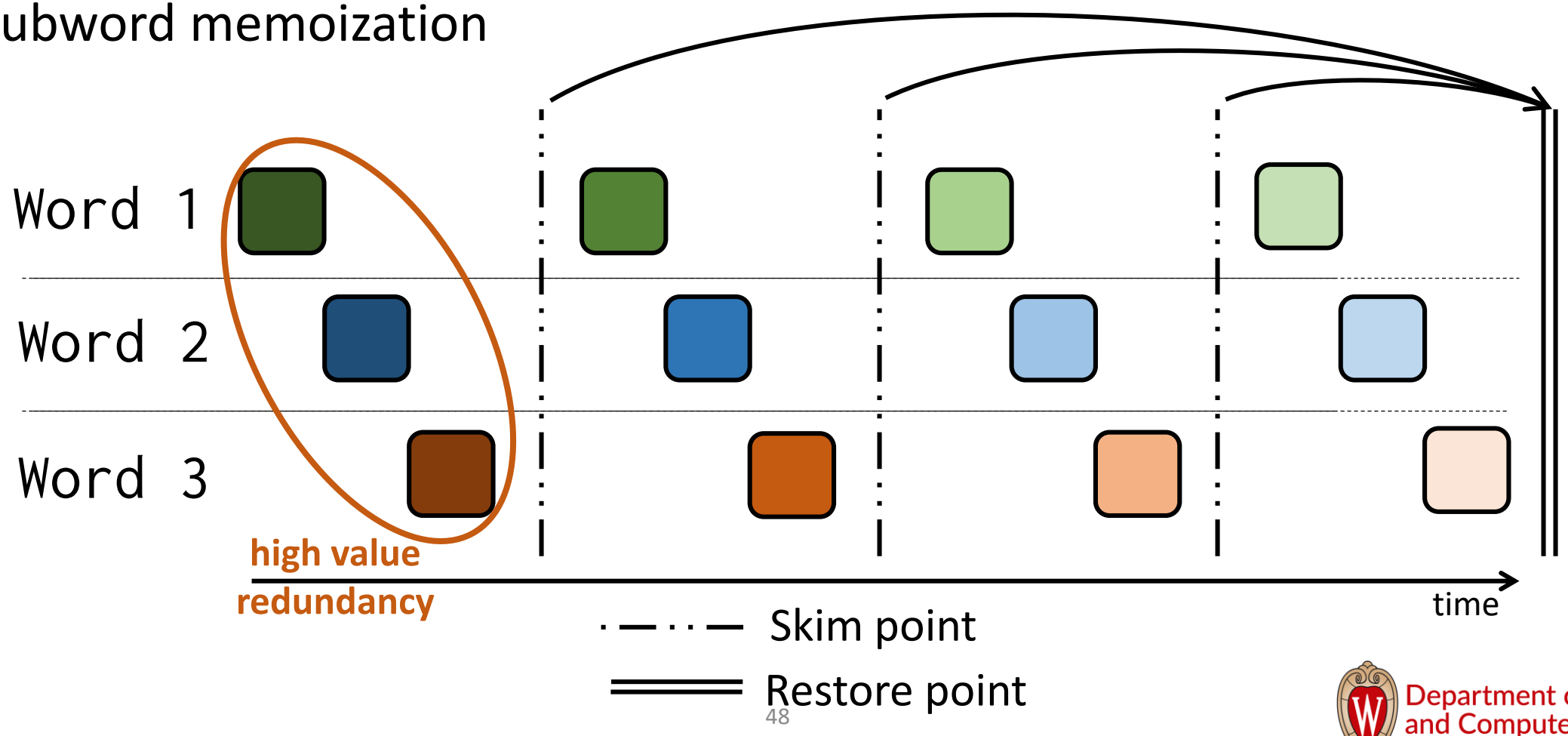


Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming – Subword Memoization

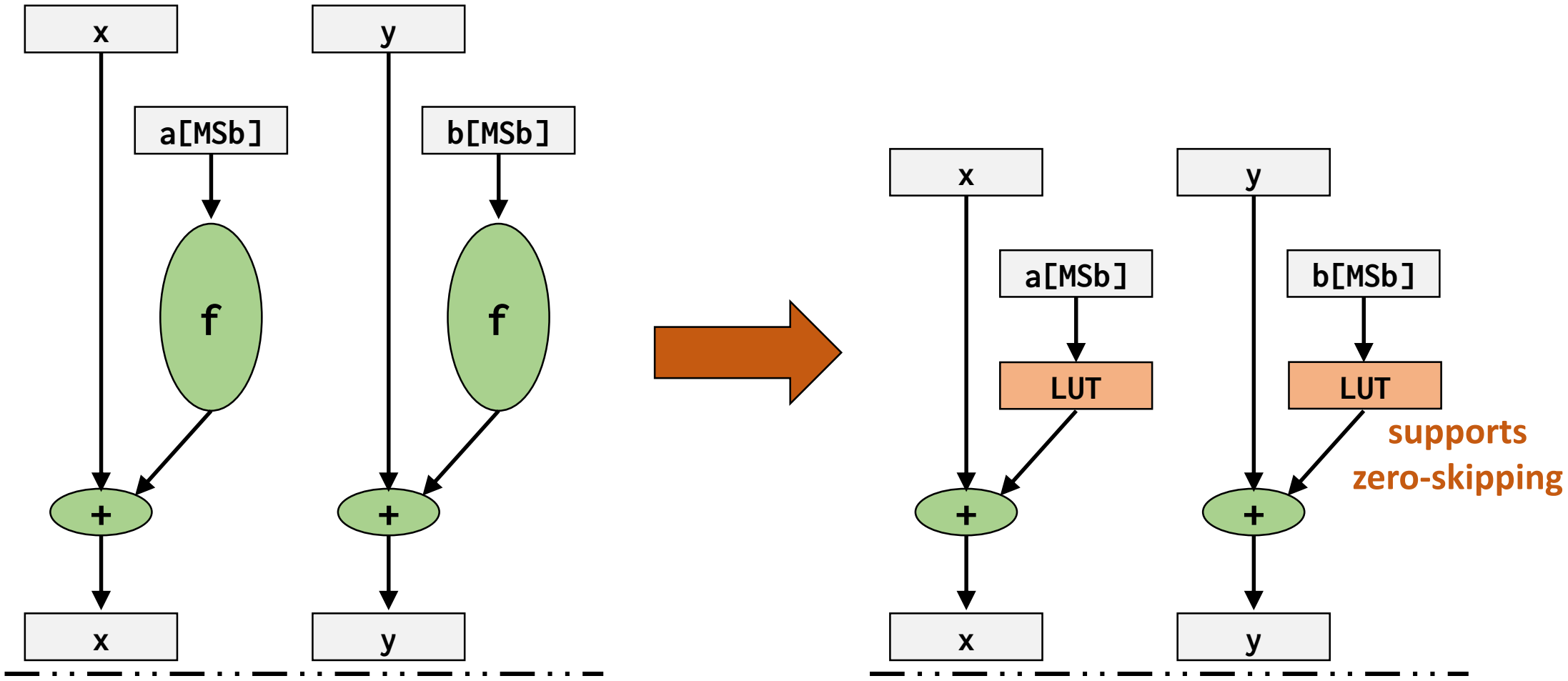
In conjunction with anytime subword pipelining:

- Subword memoization



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

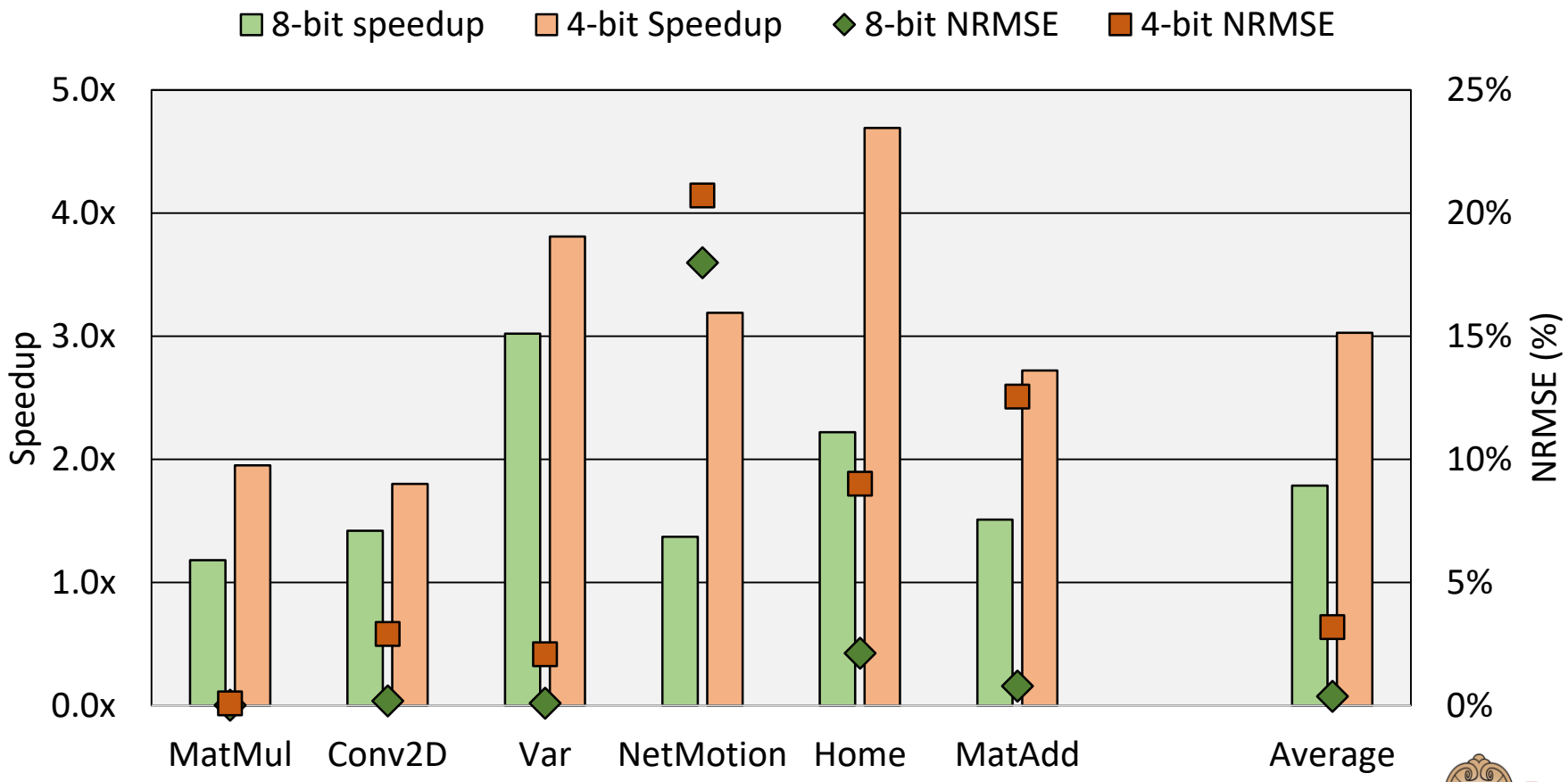
Computational Skimming – Subword Memoization



Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Computational Skimming

What's Next multi-backup system:



Conclusion

Design Tools:

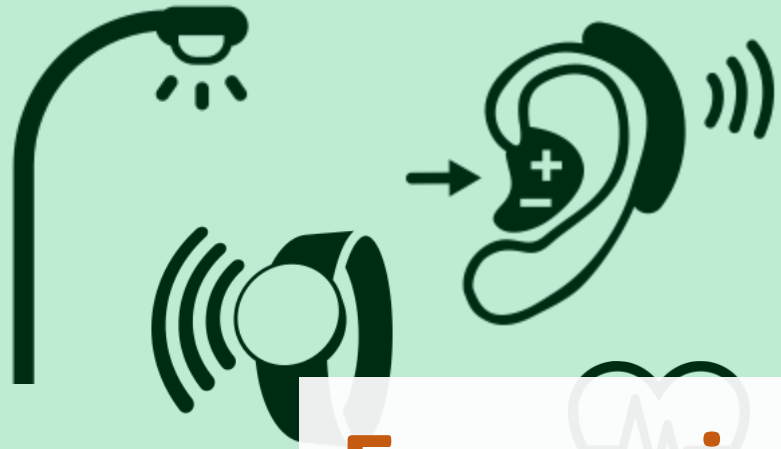
- EH Model [MICRO'18]

Design Paradigms:

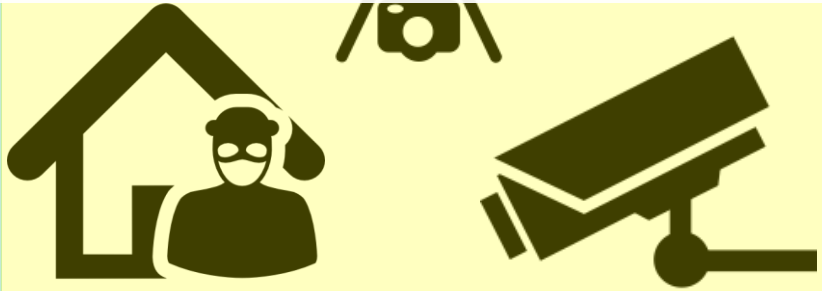
- Computational Skimming [HPCA'19]

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

Conclusion



Energy is the common denominator



long lifetime

quick response

advanced capability

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN-MADISON

Acknowledgements

Research Group:

- Abhishek Bhattacharyya
- Asmita Pal
- Di Wu
- Giri Prasanna Mugunda Krishnan
- Mitali Soni

Collaborators:

- University of Wisconsin-Madison
- University of Toronto
- IBM Research

Thank you

Joshua San Miguel
University of Wisconsin-Madison
jsanmiguel@wisc.edu

Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON