# Generating Realistic Wear Distributions for SSDs

**Ziyang Jiao**, Bryan S. Kim

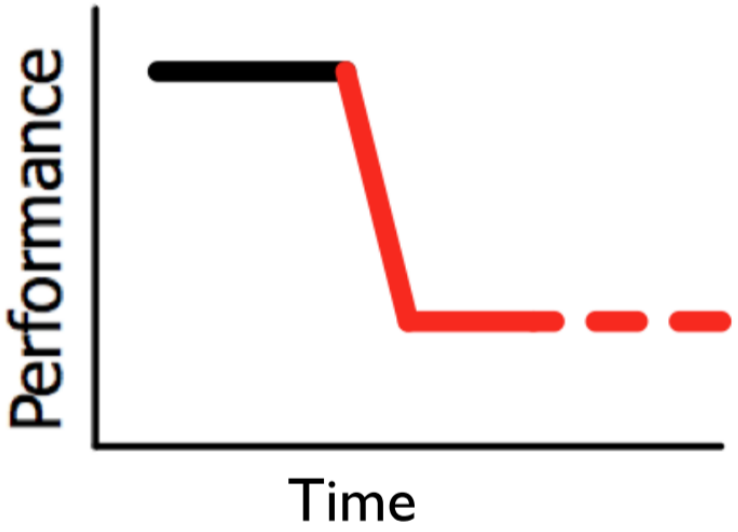
Syracuse University

# In a nutshell

- The rise of SSDs
- SSD development platforms
  - FlashSim - SIMUL 2009
  - SSDSim - ICS 2011
  - FTLSim - SYSTOR 2012
  - FEMU - FAST 2018
  - MQSim - FAST 2018
  - Amber - MICRO 2018

# In a nutshell

- The rise of SSDs
- SSD development platforms
- The fail-slow symptom of SSDs




- Haryadi S. Gunawi et al, “Fail-Slow at Scale: Evidence of Hardware Performance Faults in Large Production Systems”, FAST 2018

# Current art

*Preconditioning*:

The process of writing data to the device to prepare it for steady state measurement.

- https://www.snia.org/sites/default/files/technical-work/pts/release/SNIA-SSS-PTS-Enterprise-v1.1.pdf

# Current art

*Preconditioning*:

The process of writing data to the device to prepare it for steady state measurement.

**Expensive:**

- ✖ **Resources**
- ✖ **Time**

- https://www.snia.org/sites/default/files/technical-work/pts/release/SNIA-SSS-PTS-Enterprise-v1.1.pdf

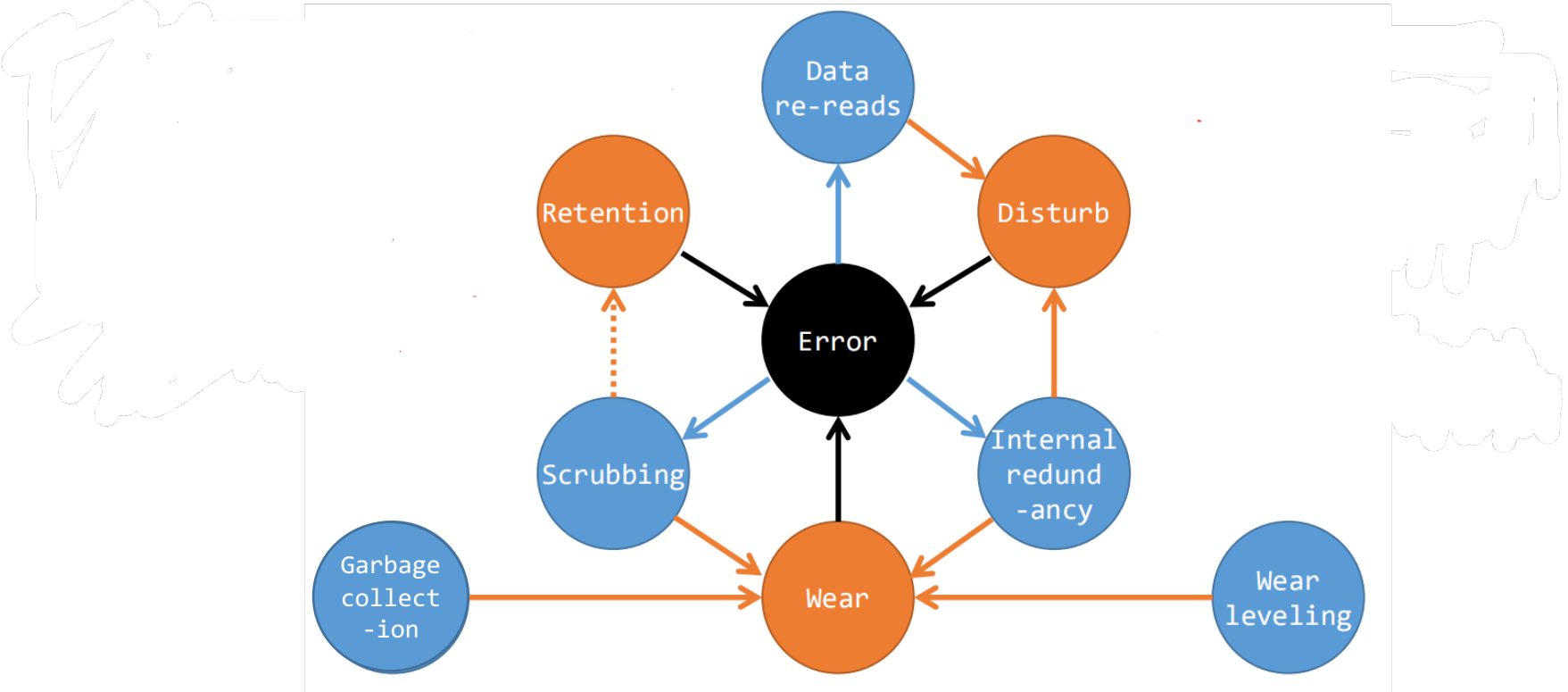
# Our approach

Fast-Forwardable SSD:

a machine learning-based SSD aging framework that generates representative future wear-out states

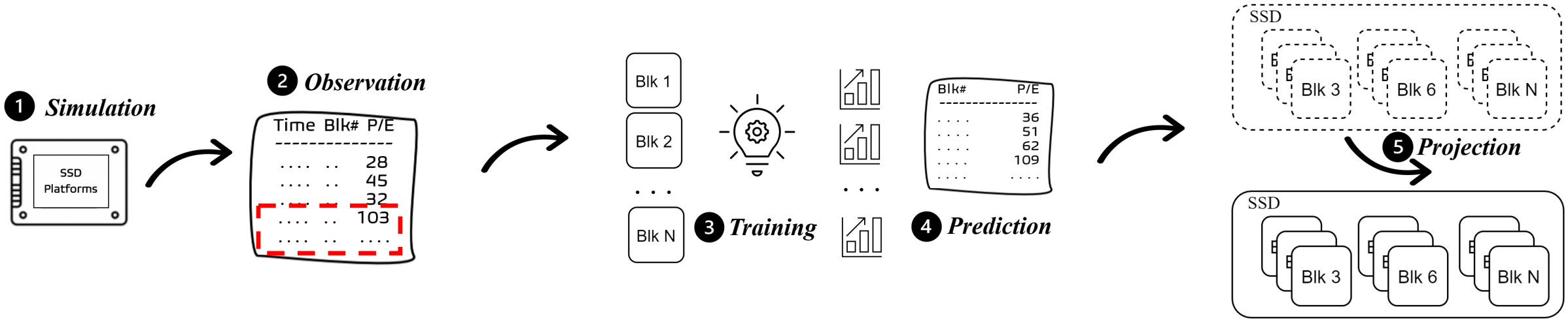
# Challenges

- The cost of machine learning models → Efficiency
- The internal intricacy of SSDs → Accuracy

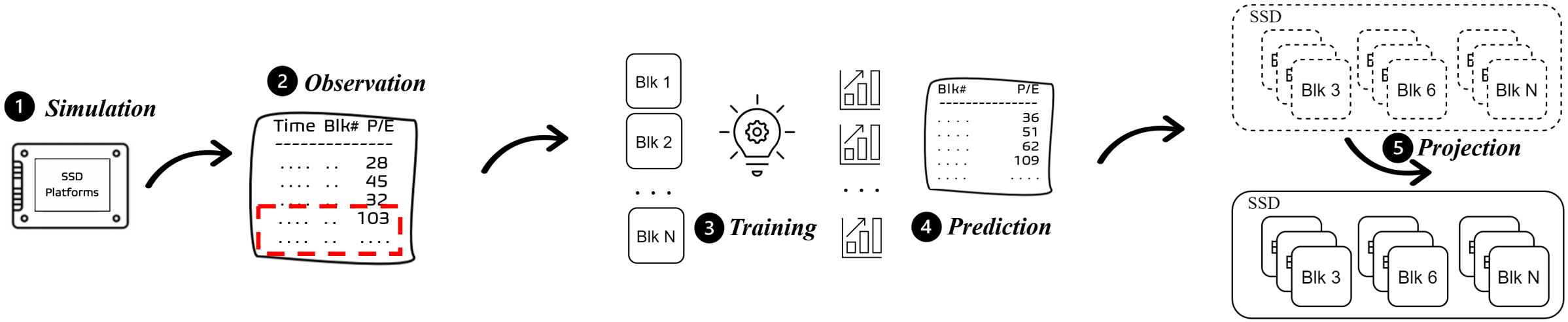



- Bryan S. Kim et al, "CPR for SSDs", HotOS 2019

# System overview



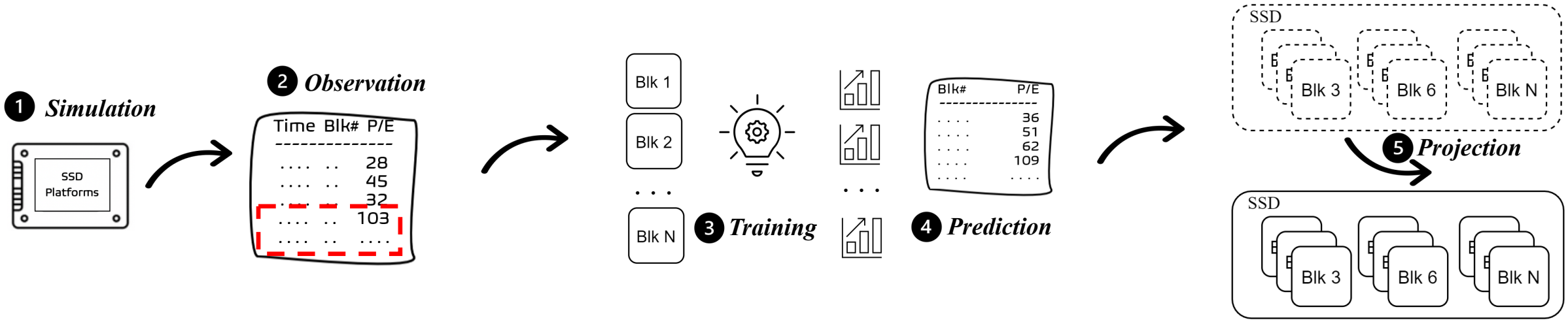
# System overview




Simulation Stage

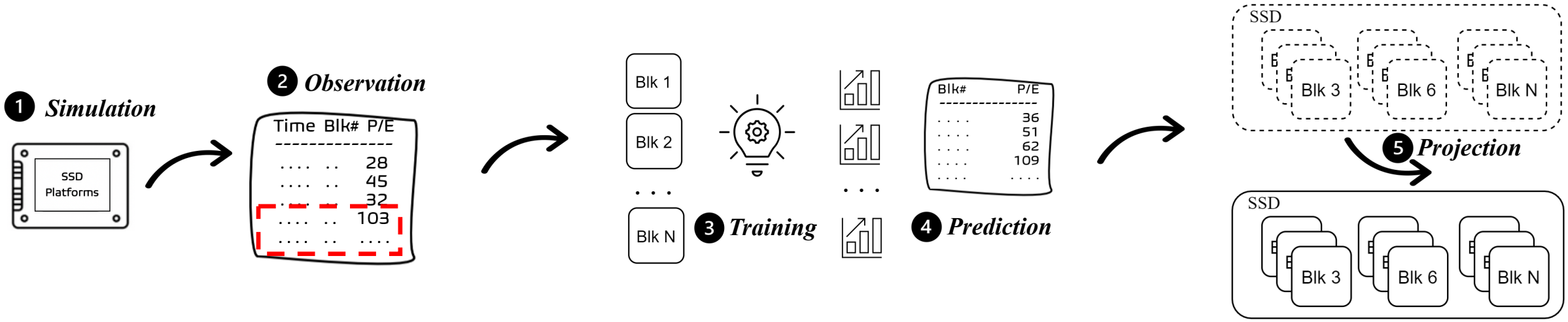
- Observe SSD's internal activities
- Output log periodically

# System overview



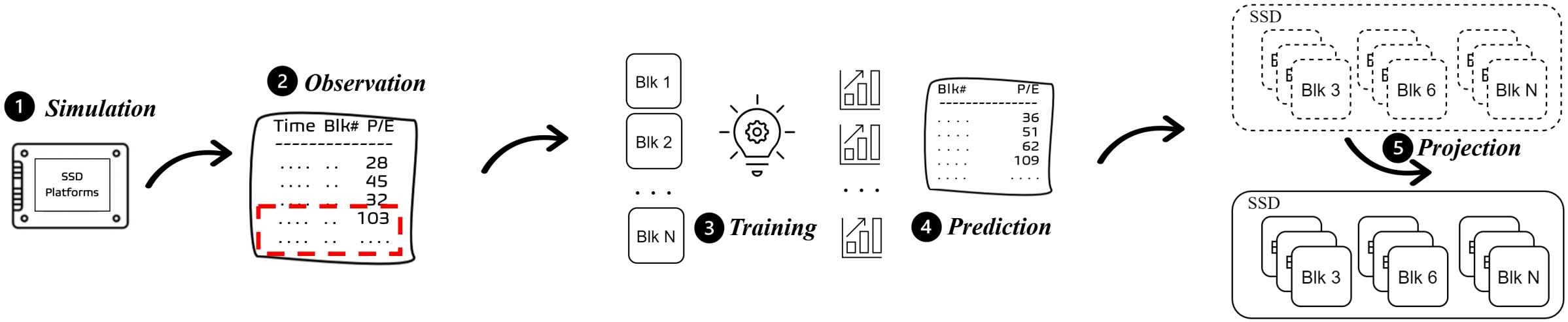

| Simulation Stage | Prediction Stage |
|---|---|
| • Observe SSD's internal activities<br>• Output log periodically | • Predict future state based on the latest history<br>• Each block has its own model<br>• Lightweight regression model (LR) |

# System overview




| Simulation Stage | Prediction Stage | Transition Stage |
|---|---|---|
| • Observe SSD's internal activities<br>• Output log periodically | • Predict future state based on the latest history<br>• Each block has its own model<br>• Lightweight regression model (LR) | • Project SSD states based on the prediction results |

# System overview




**Simulation Stage**

- Observe SSD's internal activities
- Output log periodically

**Prediction Stage**

- Predict future state based on the latest history
- Each block has its own model
- Lightweight regression model (LR)

**Transition Stage**

- Project SSD states based on the prediction results

# Prior work

Simulation acceleration with ML:

- DEVS

  DEVS execution acceleration with machine learning.
  *SpringSim 2016:* *https://dl.acm.org/doi/10.5555/2975389.2975399*

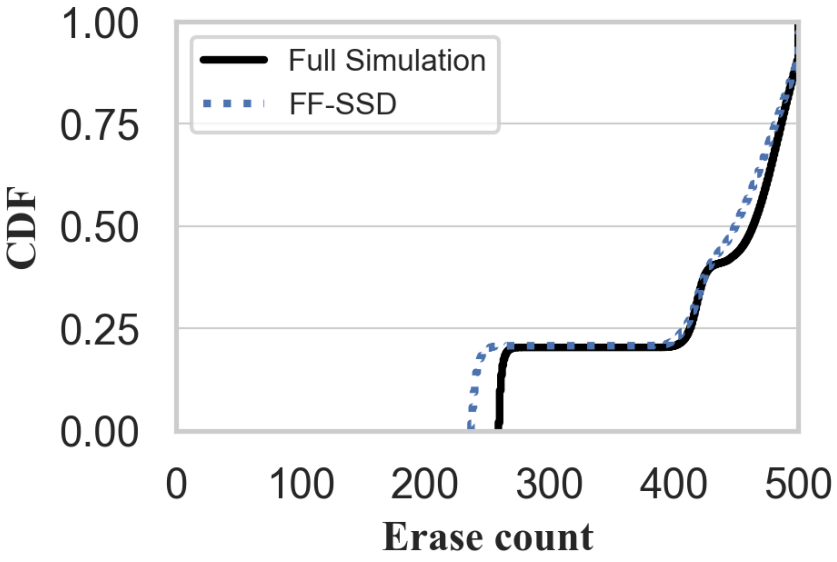
  - Consider multiple model candidates

- CML

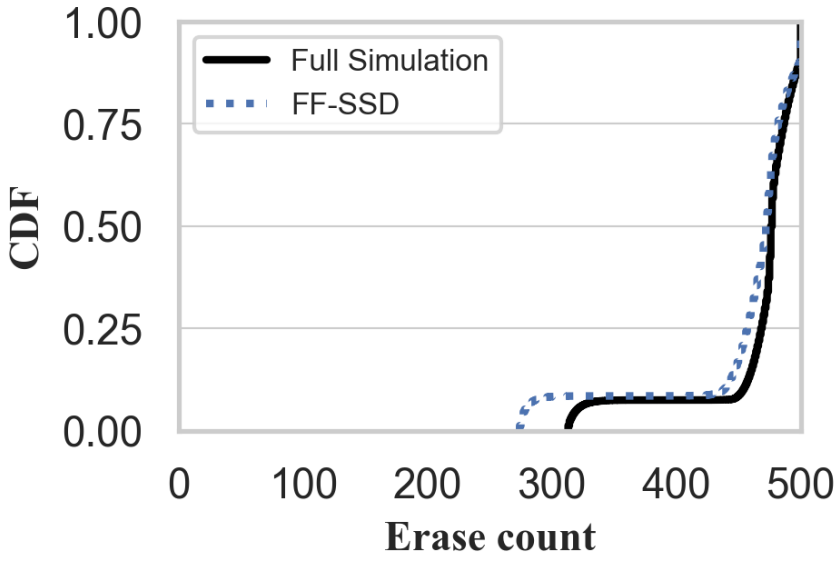
  Using continuous statistical machine learning to enable high-speed performance prediction in hybrid instruction-/cycle-accurate instruction set simulators.
  *CODES+ISSS 2009:* *https://dl.acm.org/doi/10.1145/1629435.1629478*

  - Continuously incorporate the latest data to update model

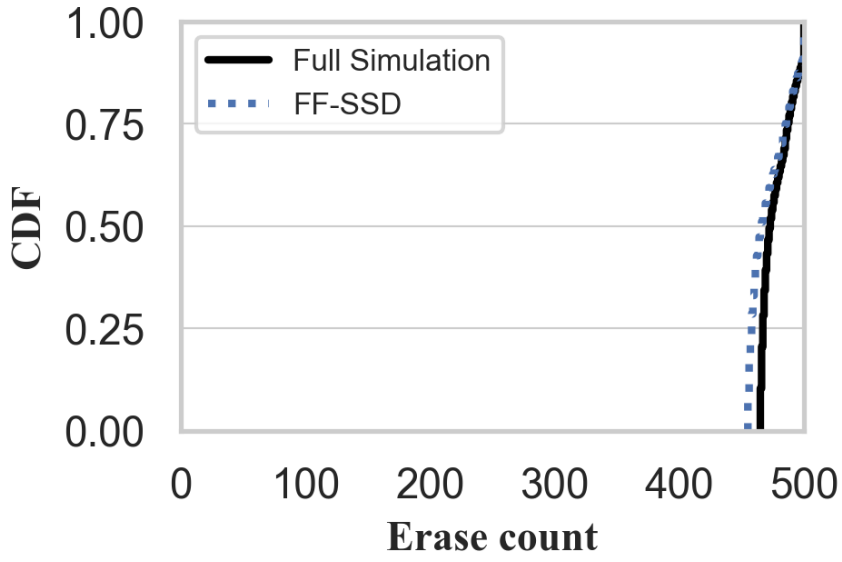
# Evaluations




Windows Build Server



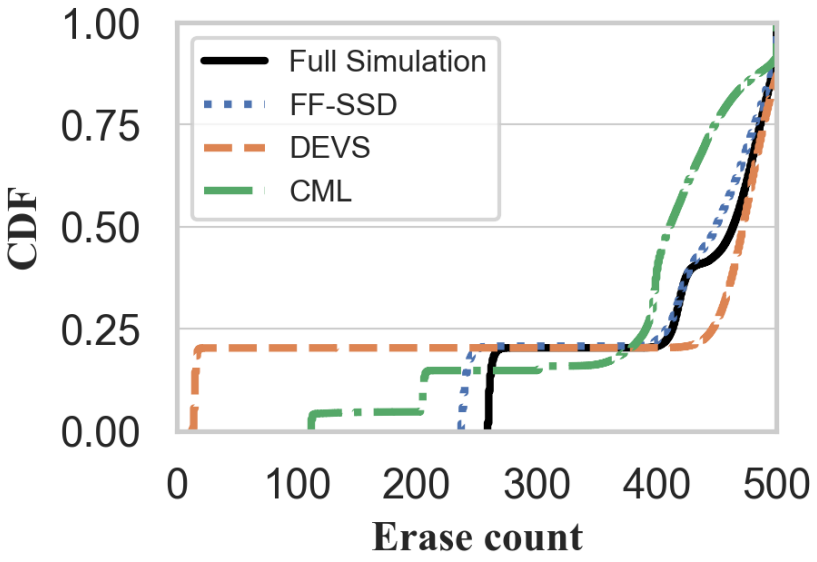

Developer Tools Release



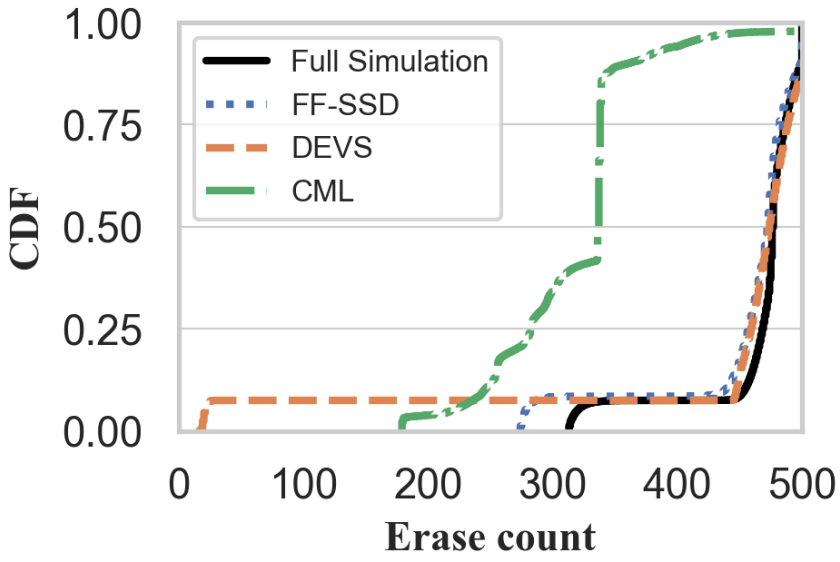

Virtual Desktop Infra

Evaluation under real-world workloads. With one-third workload saving compared to full simulations, FF-SSD generates the final states of SSD with high accuracy.

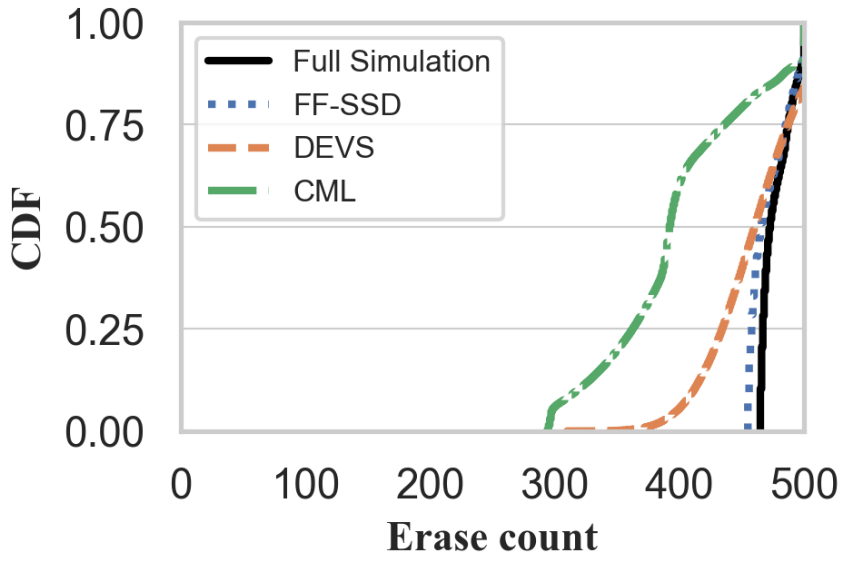
# Evaluations




Windows Build Server




Developer Tools Release




Virtual Desktop Infra

Evaluation under real-world workloads. With one-third workload saving compared to full simulations, FF-SSD generates the final states of SSD with high accuracy.

# Future work

- Analyze *FF-SSD* over various workloads
- Expand *FF-SSD* to support different platforms and SSD models
  - Platforms:
    - Amber
    - FEMU
    - FTLSim
    - …
- Improve both prediction accuracy and aging efficiency further

# *Thank you!*

*Ziyang Jiao*
*zjiao04@syr.edu*