Extracting Flexible, Replayable Models from Large Block Traces

STONY BROOK UNIVERSITY — IBM — HARVEY MUDD COLLEGE

V. Tarasov⋆, S. Kumar⋆, J. Ma †, D. Hildenbrand×, A. Povzner×, G. Kuenning †, E. Zadok⋆

⋆ Stony Brook University
† Harvey Mudd College
× IBM Research, Almaden

Production I/O traces are valuable as they embody real-world workloads. Unfortunately, traces are cumbersome to replay. **Benchmarks**, on the other hand, are easy to setup and use, but lack the realism of production workloads. In this project we bridge the gap between benchmarks and traces by developing a methodology and a set of tools for **automatic conversion** of large I/O traces into concise workload models that can later be used as benchmarks.
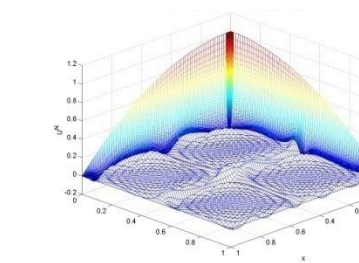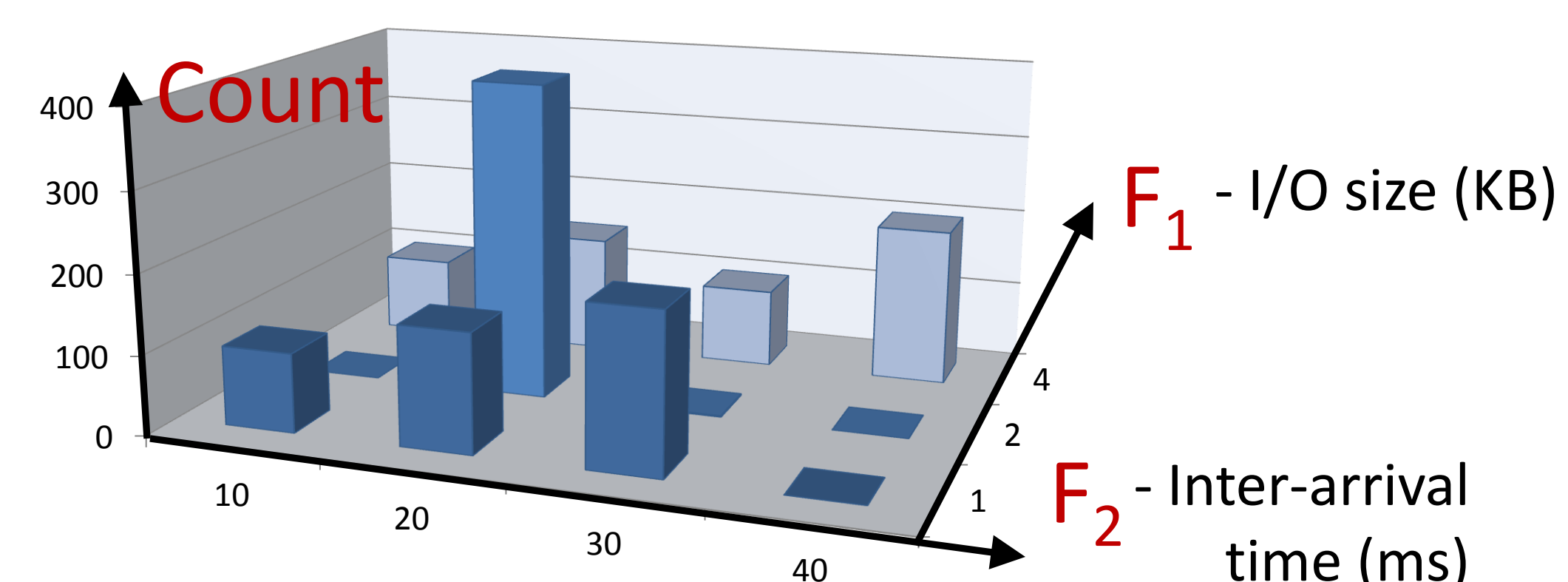
❶ Traces
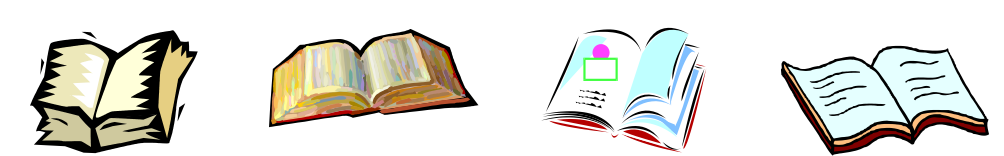
- Represent real-world workloads 😊
- Credible: often used in the research
  35 out of 68 papers in FAST, ATC, OSDI, HotStorage, and MSST in 2010
- Quite a few are collected
  SNIA's trace repository growing, dozens currently available
- Massive in size
- Cumbersome to replay 😞
- Hard to scale
- Difficult to adjust even to small workload changes

❷ Benchmarks

- Great for micro-benchmarking 😊
- Easy to set up and run
- Usually scale well
- Workloads are easily modifiable 😞
- Workloads generated might be not relevant to the real world
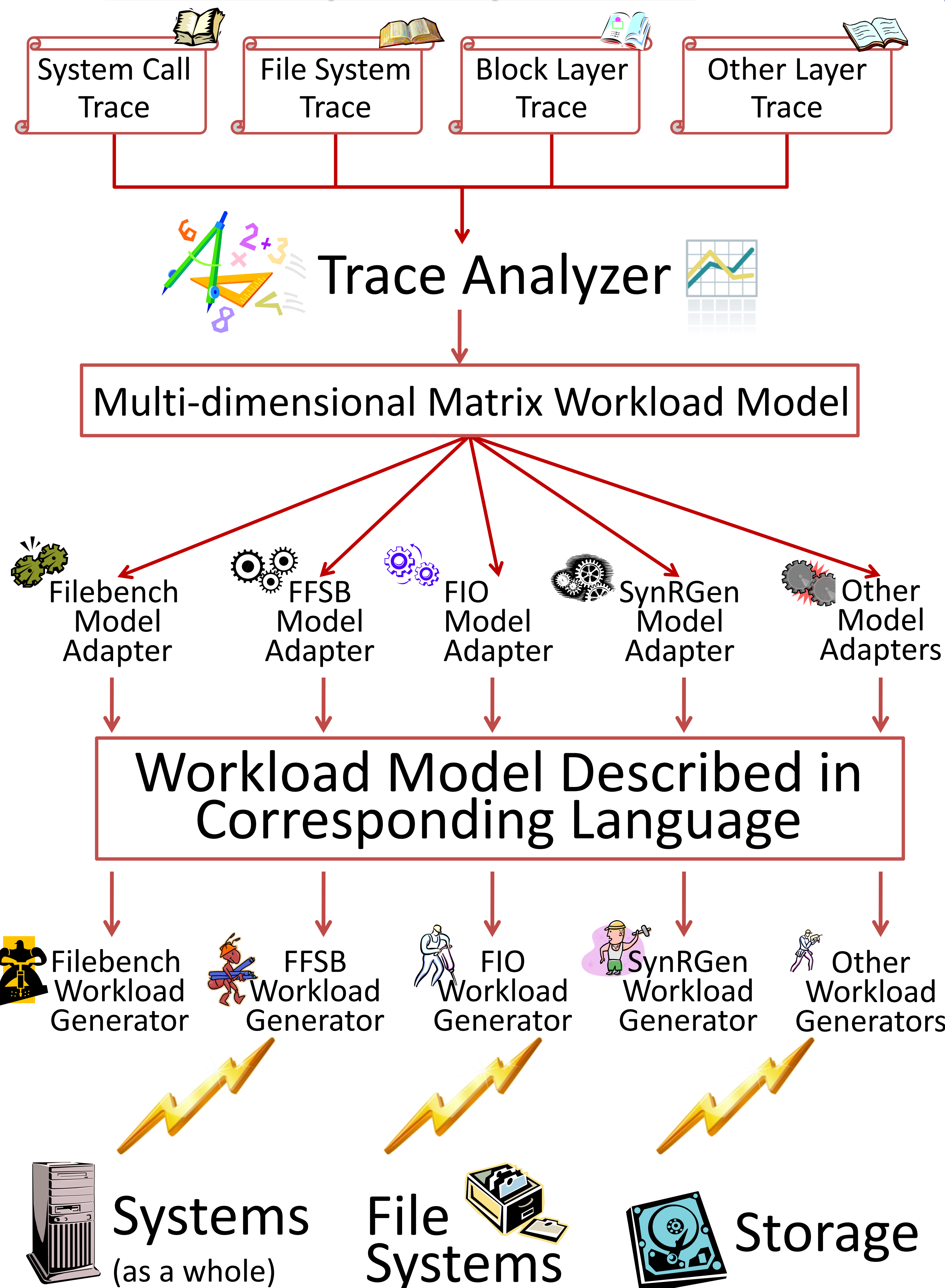
Can we combine the best of two worlds?

❸ Trace chunking

Average behavior over **entire** trace loses time variance ➡ **Divide** trace into self-consistent chunks

0 —— chunk 1 — chunk 3 — … — chunk N —— T
chunk 2 … …
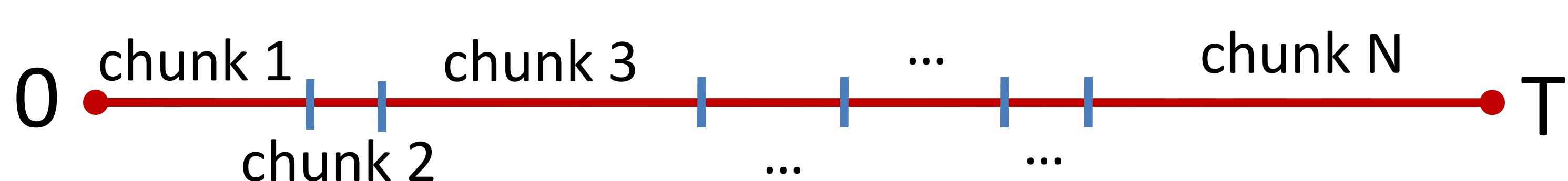
- **Fixed** chunking as a first step
- Perform chunk **deduplication** to save trace space:
  - Identify chunks, where workload is similar
  - Store only **single instance** of similar chunks
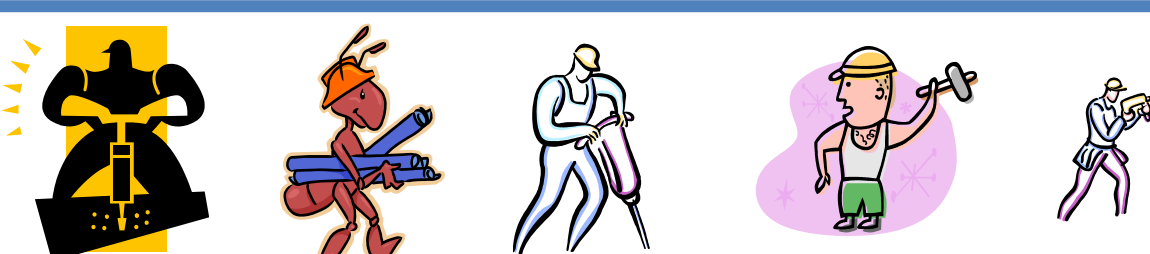 - Maintain a **mapping** between trace time and unique chunk instances

❹ Modeling a chunk

- The number of **workload parameters** is **infinite**
- Select workload parameters:
  - Parameters **used** in the studies **before**
    ✓ Read/Write ratio, Inter-arrival time, I/O size, …
  - Make the tool **extensible**
- Use **multi-dimensional** histogram of **feature functions**

Count — 400, 300, 200, 100, 0
F_1 - I/O size (KB)
F_2 - Inter-arrival time (ms)
10 20 30 40

❺ Complete picture

System Call Trace | File System Trace | Block Layer Trace | Other Layer Trace

→ Trace Analyzer

→ Multi-dimensional Matrix Workload Model

→ Filebench Model Adapter | FFSB Model Adapter | FIO Model Adapter | SynRGen Model Adapter | Other Model Adapters

→ Workload Model Described in Corresponding Language

→ Filebench Workload Generator | FFSB Workload Generator | FIO Workload Generator | SynRGen Workload Generator | Other Workload Generators

→ Systems (as a whole) | File Systems | Storage