Flexible Data Placement (FDP): Use Cases

Rory Bolt
KIOXIA America, Inc.
FARC-203-1
August 6th, 2025

FMS
the Future of Memory and Storage

KIOXIA

Best Practices From A Device Perspective

- **If you want to approach a Write Amplification Factor (WAF) of 1.0:**
  - Overprovisioning helps (at a cost, of course). This can include “short stroking” in addition to the manufacturer’s factory-configured overprovisioning. This minimizes unnecessary write amplification due to race conditions
- **Assuming direct logical block addressing (LBA) access:**
  - Log structured write algorithms and circular buffers generally work well with FDP
    - Log Structured Merge (LSM) of RocksDB[1] is a great example
    - Transaction journals
    - Copy-on-write snapshotting file systems
  - Separate hot and cold data whenever possible
    - Traditionally classifying data at creation is very difficult
    - Copy-on-write storage systems that perform application-level garbage collection can do this naturally during garbage collect
- **If you can’t afford to fundamentally change your storage applications, look for low hanging fruit**
  - The Pareto Principle[2] usually holds true; just separating out a small portion of your data may yield large results (i.e. metadata is often far more volatile than user data, temp files can be moved to a dedicated namespace with a default placement ID)

1. RocksDB is a database engine optimized for high read/write throughput performance and low latency storage.
2. The Pareto Principle, also known as the 80/20 rule, is a concept stating that roughly 80% of effects come from 20% of causes.

KIOXIA RocksDB Plugin for FDP “UFROP”

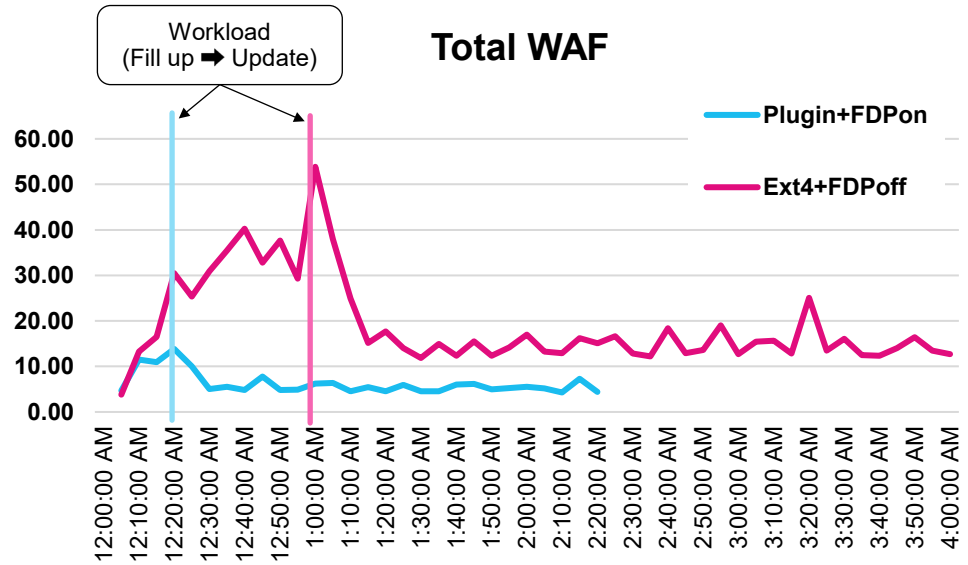
Evaluation Result – WAF

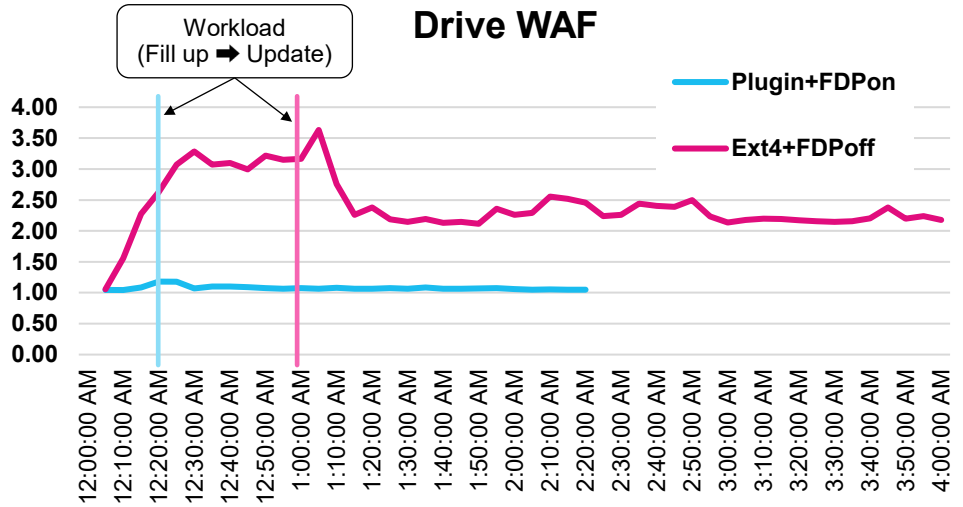
- **WAF comparison (average for 1h at steady state)**

Back-end	FDP	Drive WAF*	Total WAF
Plugin	**On**	1.06	**5.34**
Ext4	**Off**	2.20	**15.01**

*Including additional writes of SSD internal information.

- **Conditions**
  - KIOXIA XD8 Series SSD (8 TB[1] with special firmware to shrink capacity to 256 GB)
  - Workload: Drive filling up to 67% → Updates
  - Key : Value = 20B : 16KB

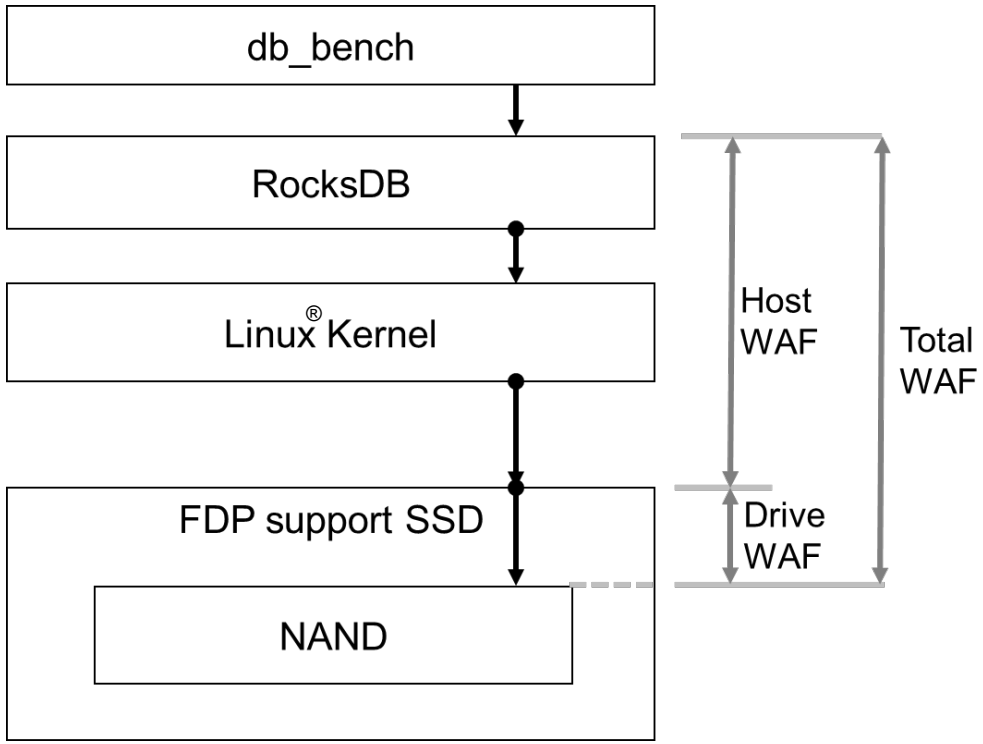


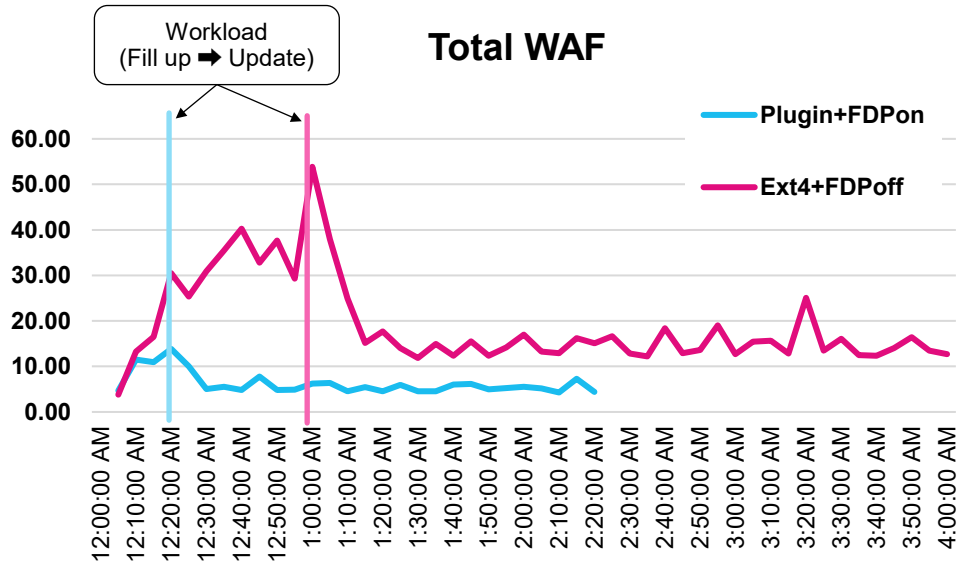


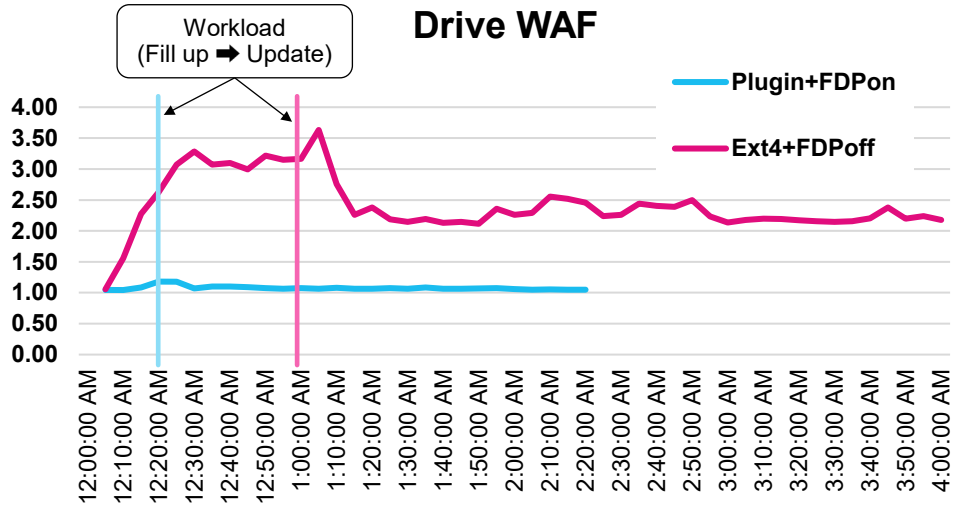
Source: Tables and charts created by KIOXIA engineers. Images and/or graphics within this slide are the property of KIOXIA America, Inc. (KIOXIA) and are reproduced with the permission of KIOXIA. 1. Definition of capacity: KIOXIA Corporation defines a megabyte (MB) as 1,000,000 bytes, a gigabyte (GB) as 1,000,000,000 bytes, a terabyte (TB) as 1,000,000,000,000 bytes and a petabyte (PB) as 1,000,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1Gbit = 2^30 bits = 1,073,741,824 bits, 1GB = 2^30 bytes = 1,073,741,824 bytes, 1TB = 2^40 bytes = 1,099,511,627,776 bytes and 1PB = 2^40 bytes = 1,125,899,906,842,624 bytes and therefore shows less storage capacity. Available storage capacity (including examples of various media files) will vary based on file size, formatting, settings, software and operating system, and/or pre-installed software applications, or media content. Actual formatted capacity may vary.

Evaluation Result – WAF

Measurement Points



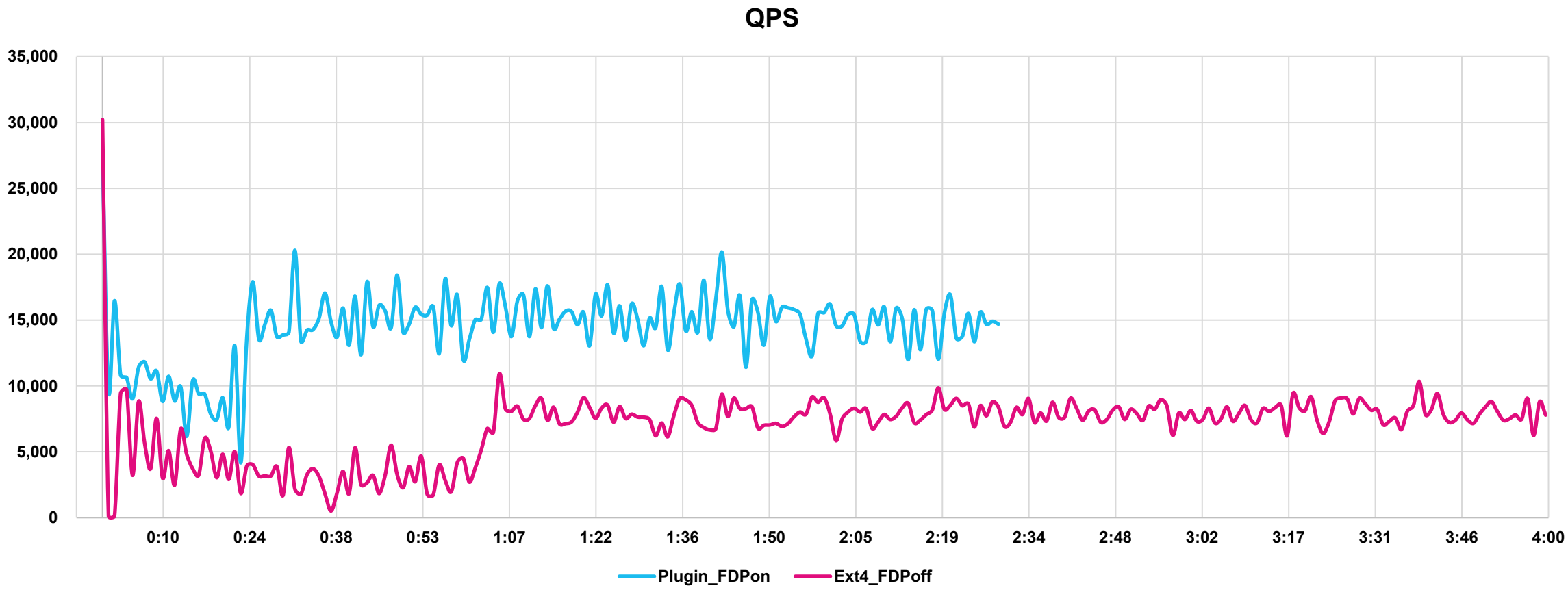




Source: Tables and charts created by KIOXIA engineers. Images and/or graphics within this slide are the property of KIOXIA America, Inc. (KIOXIA) and are reproduced with the permission of KIOXIA. Linux is a registered trademark of Linus Torvalds and its use is governed by the Linux Trademark Institute.

Evaluation Result – Throughput

Backend	FDP	Throughput [KQPS[1]]
Plugin	On	15.0
Ext4	Off	8.0



OSS Contribution & Further Work

- **The plugin has been released as open-source software**
  - **https://github.com/kioxiaamerica/ufrop**
  - **Efforts underway to include in list of 3rd party RocksDB Plug-ins of main repository**

- **Planning to update for:**
  - **Multi-drive support (striping)**
    - **(Optional) Error correction functionality**
  - **Further WAF improvements**

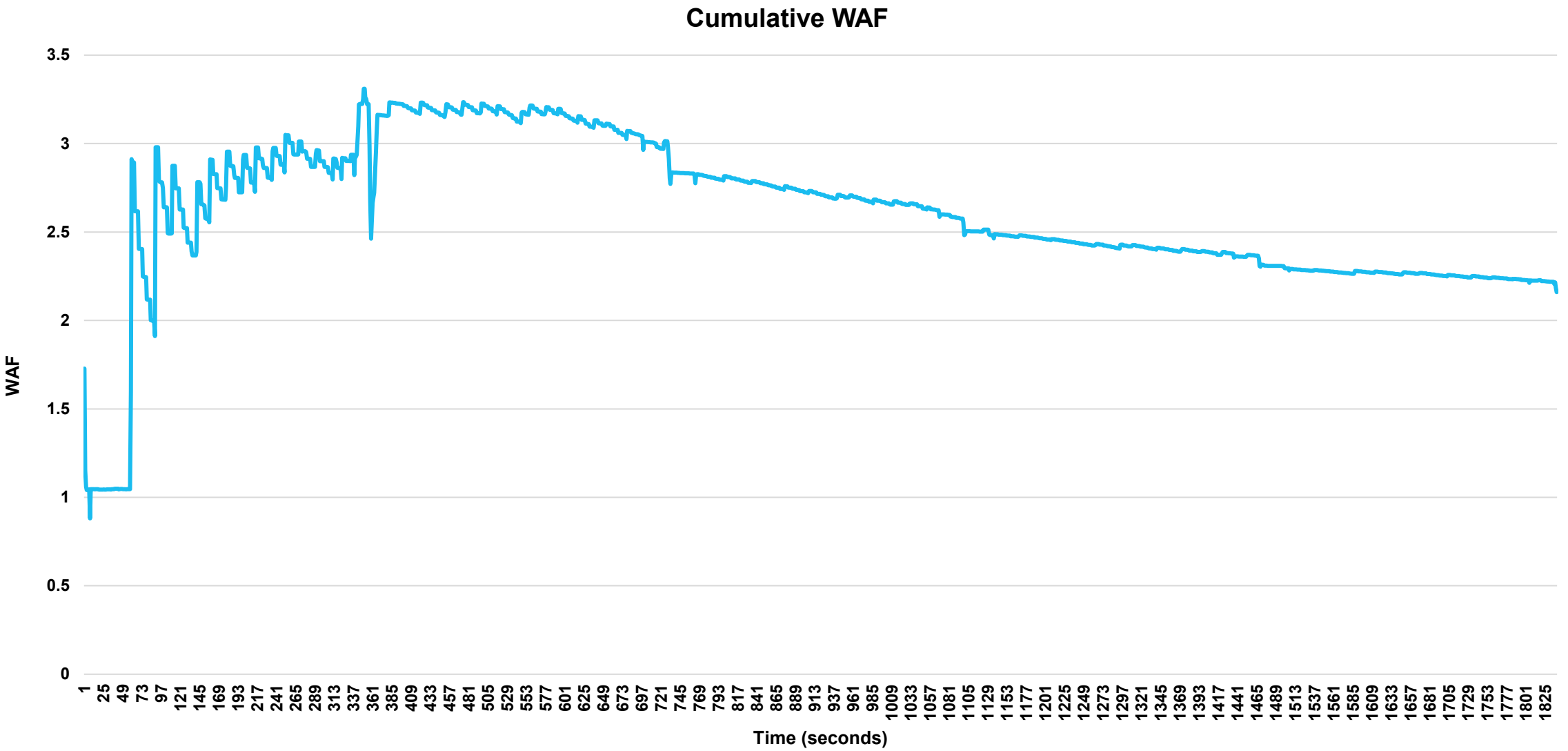
Flexible Data Placement Low Hanging Fruit: *Results From a Real Workload*

A Real World Experiment

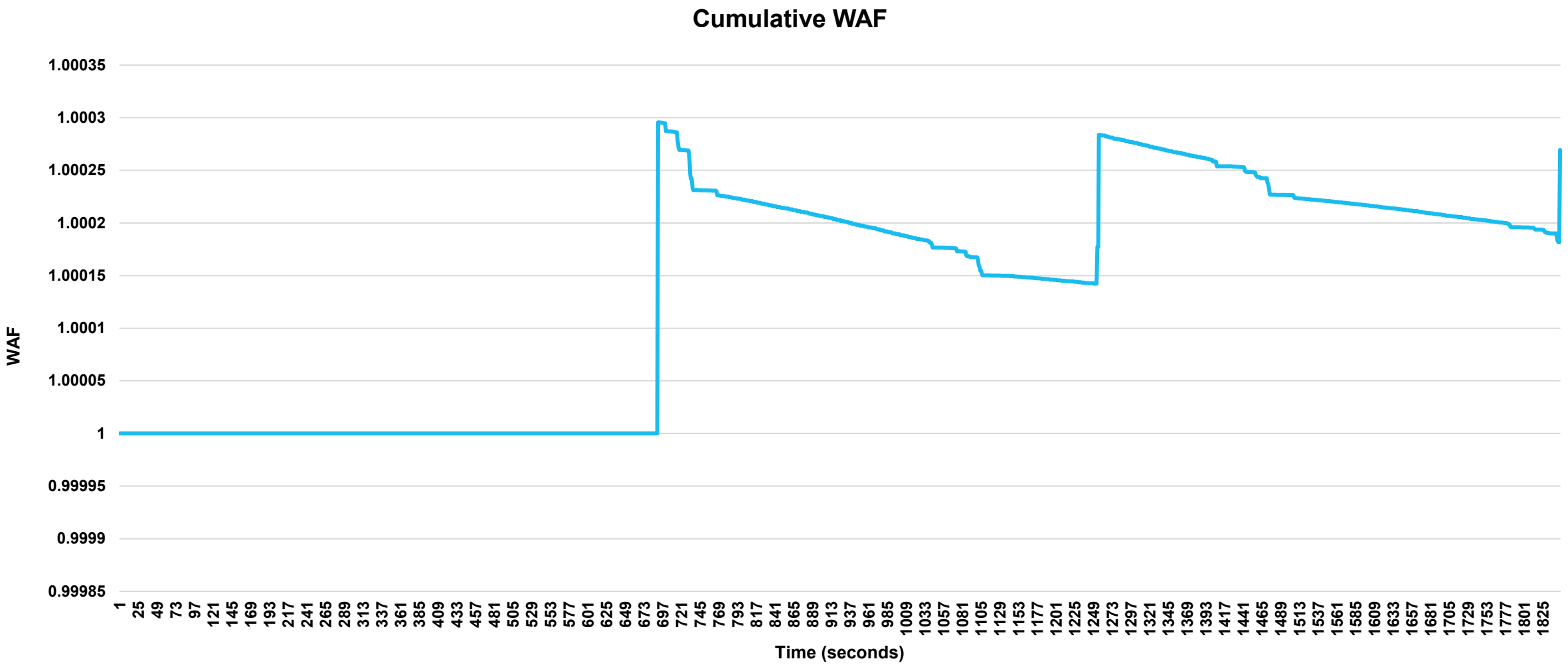
Theory: Temporary files and swap space are used for data with a short lifespan. By separating this data out from data with a longer lifespan we should see a reduction in WAF.

- The exact same workload run on the same machine with FDP disabled and enabled
- Separate namespace used for the /tmp filesystem with a different default placement ID specified
- A real workload of compiling the Linux® kernel was selected and a Continuous Integration/Continuous Deployment (CI/CD) environment was modeled. Compilers often generate temporary files.
- No software changes were made

Non-FDP Results



FDP Results



Note the scale! This is a fantastic result of almost 1.0 WAF

Source: Tables and charts created by KIOXIA engineers. Images and/or graphics within this slide are the property of KIOXIA America, Inc. (KIOXIA) and are reproduced with the permission of KIOXIA.

A simple configuration change yielded large results on a real workload

- WAF was reduced from about 2.2 to almost exactly 1.0
- This would effectively double the lifespan of a device used for this workload
- This would reduce the energy consumed by the device

KIOXIA