AUTOMOTIVE (→) WORKLOAD ANALYSIS

"WHITE PAPER"

APRIL.2026

AUTOMOTIVE (→) WORKLOAD ANALYSIS

Overview

The automotive industry is going through a major transformation driven by the adoption of new technologies, applications, and services. In the evolution to the next-gen automotive system, storage demand is growing and evolving to support more demanding usage cases and new applications. Advanced infotainment, ADAS systems, data-recording sensors and cameras, and 3D mapping are all expected to generate and consume huge amounts of data. Data needs to be kept in the vehicle storage to support real-time decisions. Entirely new applications and capabilities are expected to arise, utilizing 5G capabilities, artificial intelligence, machine and deep learning, and big data analysis.

As applications change and workloads increase, the storage requirements need to be reevaluated. While smartphones' write workload has steadily increased over time, as presented in Figure 1, the automotive industry is expected to see an incremental increase in storage usage resulting from autonomous-drive complexity and regulation requirements.

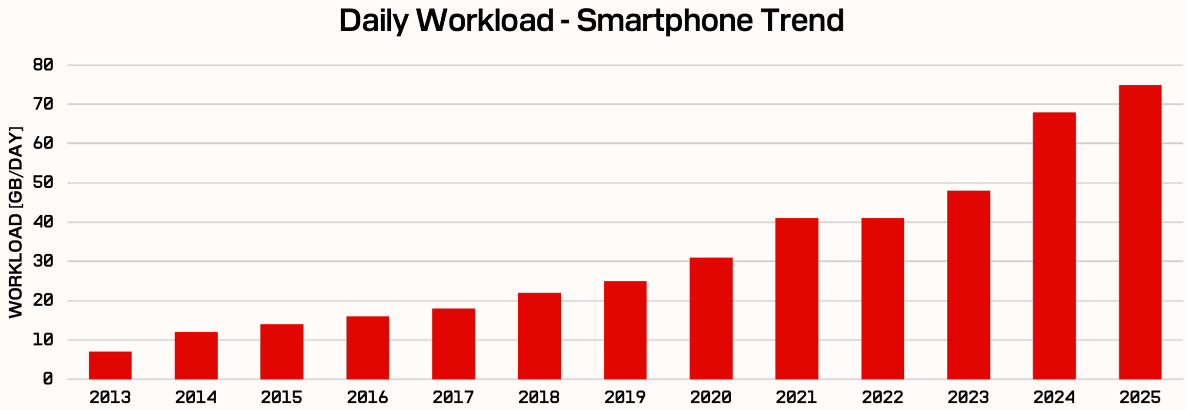

Figure 1. Sandisk estimated daily storage workload of an extreme smartphone user[1]

Storage is a critical component in the ongoing evolution of the automotive industry, yet it's not just about having the capacity to store the growing amount of data. It's about having the ability to support the durability and longevity that automotive systems require, supporting features that allow monitoring and providing the best user experience possible to address the new applications' needs.

Sandisk has developed tools and analysis methodologies to architect the right solutions for current and future unique workloads to support the growing automotive storage market and prevent expensive recalls due to over-written storage.

The Importance of Workload Analysis

Real-life workload analysis is highly important to understanding the automotive target system needs and requirements from the storage. Storage workload analysis enables a proactive monitoring and prevention of software issues and helps customers determine what capacity is needed to prevent storage wear-out. It is a baseline for lifetime prediction, storage product design and validation considerations. On top of it all, it can also serve as the basis for future storage and system optimizations and innovations.

The following are just a few examples of questions that may be answered using workload analysis tools:

- How much data is read and written per day in a target system?
- Are there any unexpected or unique data patterns in the target system?
- What are the data transfers and what is the implication on the write amplification factor?
- What does the storage operating profile look like in the system?
- Is the host software stack efficient?
- What are the system performance requirements?
- What is the expected lifetime of the storage under this system usage?

Over the past 10 years, Sandisk has invested and developed state-of-the-art tools and methodologies to support workload analysis activities to answer the above questions and more.

1. Data in graph is based on Sandisk internal user profiles which capture popular applications and intensive usage cases. Data is collected from Sandisk internal lab testing while examining these user profiles.

The Workload Analysis Flow

Workload analysis is composed of four main phases:

- 1st phase – bring up the infrastructure
- 2nd phase – workload definition
- 3rd phase – workload experiment
- 4th phase – system analysis

In the 1st phase (bring up the infrastructure), the main goal is to enable a tracer that will capture all desired storage events on the target system. For full analysis capabilities, it is highly recommended that the traces will include all storage transaction types, including read, write, unmap/erase and sync commands as well as management and query commands. It is also preferred that the traces will include timestamps for start and completion of each transaction/command.

There are several types of tracers that can be enabled, either from the host software or from the storage device. Some tracer types, such as eMMC tracer ftrace, UFS Ftrace SCSI tracer, and UPIU tracer, are already available as part of Linux® open source and can be enabled via Linux standard debug directory commands.

For more information and support on storage tracer enablement within your target system, please contact your local Sandisk representative.

Understanding the context of the overall usage case and finding the correlation between user actions and storage activities is a key aspect of a comprehensive workload analysis. Therefore, in the 2nd phase (workload definition), the main goal is to define the profile of user activities on the target system.

As part of the workload definition phase, questions like the below should be answered:

- What is the product/system we will analyze? (Infotainment? Data recorder? Mapping?)
- What is the storage device precondition? (Fill-up state? Which file types?)
- How many hours of usage per day?
- What types of applications represent the actual usage?
- Which user actions are performed in each application and at what cadence?
- Does this workload definition represent typical, heavy or extreme usage of the system?

Once the system is ready with the required applications and enabled tracer, the 3rd phase (actual workload experiment) can begin. During the experiment itself, huge trace logs are captured while running the defined workload. For the workload analysis to be most effective, it is highly recommended that test logs are long enough to capture the events of at least a few hours or a full day of usage.

In the 4th phase (system analysis), the captured workload logs go through post-processing and are analyzed for insights and conclusions around system and storage patterns and behaviors.

The following are some examples for storage metrics that are usually analyzed:

- I/O operations spread on the address span over time (see example in Figure 2)
- Chunk sizes distribution
- Sequential versus random operations
- Storage activity profile over time
- Storage IOPS and MB/s over time
- Command latencies over time
- Queue utilization (depth) over time

Sandisk can support its customers with storage trace analysis and insights.

Figure 2. Operation spread on the address span over time.

Example: Identifying Unexpected Write Activity Based on Workload Analysis

Conducting workload analysis as a repetitive routine on evolving product generations and different setups allows identification of changing trends and unexpected patterns.

Consider the below example, where the same navigation application was tested on two different setups. As this application was tested over the years in different setups and conditions, its behavior and patterns are well known and produce ~150–300MB in 20 minutes of navigation. However, while analyzing this application on a more recent setup, unexpectedly, it generated 1.5GB of data under the same exact usage profile.

Analyzing the operations on the address range (see Figure 3) showed a specific area of addresses which absorbed most of the writes. Using advanced analysis tool capabilities to zoom in on this area (see Figure 4) we uncovered intense logging activity which continuously re-wrote, in a cyclic manner, the same 1MB file.

Setup #1 20 Min. Navigation ~250MB writes

Setup #2 20 Min. Navigation ~1.5GB

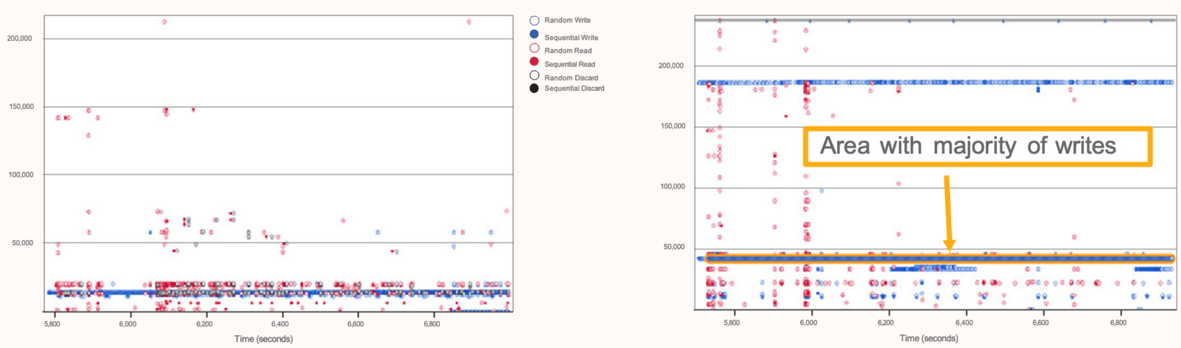

Figure 3. Excessive writes observed on setup #2 while running navigation app.

Using workload analysis methodology to identify such unexpected behavior before it is implemented and operated in the field is a critical capability that enables prevention and allows corrective actions ahead of time.

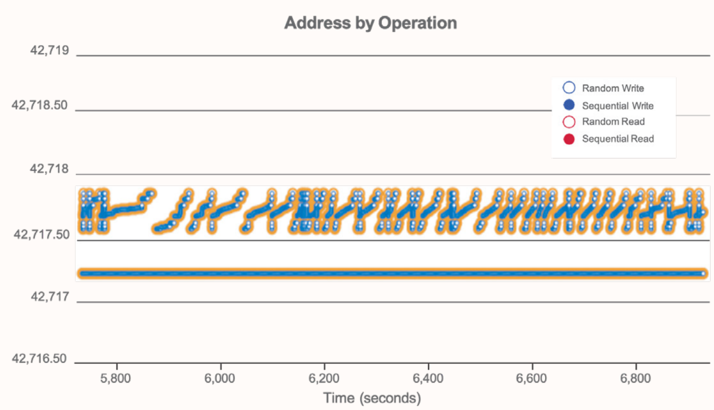

Figure 4. Typical logging activity behavior, as observed on setup #2.

Product Lifetime Prediction Based on Workload Analysis

Workload analysis is a basic tool for enabling storage lifetime prediction in the target system.

The lifetime prediction is done based on Write Amplification Factor (WAF) and TeraBytes Written (TBW).

WAF is a ratio between the amount of data written to the flash memory and the amount of data written by the host. Write Amplification (WA) is a phenomena associated with flash memory where the actual amount of physical information written to the flash is usually more than the amount intended to be written by the host. The main reasons for WA are the flash memory attributes which can only erase fixed-size blocks, cannot overwrite programmed data, and need to refresh the data in certain conditions. The WAF formula is shown in Figure 5.

The flash memory capacity and write cycles are given; therefore, lower WAF helps to achieve longer product lifetime.

$$\frac{\text{Data written to the flash memory}}{\text{Data written by the host}} = \text{Write Amplification Factor}$$

Figure 5. Write Amplification Factor formula.

WAF changes per workload pattern and product. WAF analysis is provided based on simulation or testing of a certain product under the target typical or worst-case workload. The results can be impacted by many metrics and behaviors which are determined as part of the workload analysis flow. Some examples for elements which may impact WAF are listed below:

- Host write pattern – sequential versus random, hot spots, chunk sizes
- Discard commands and sync cache usage
- Idle time profile and storage garbage-collection algorithms during that time
- Free space on the physical media
- Flash management metadata and tables updates

Product TBW is the total amount of host writes, in terabytes, that the product can support in its lifetime. TBW is directly impacted by the workload WAF. The expected lifetime prediction can be calculated based on the product TBW and the expected typical/heavy daily writes concluded from the workload analysis. See Figure 6 for the formulas to calculate both TBW and expected lifetime.

$$\text{TBW} = \frac{[\text{Exported Capacity}] \times [\text{Max Program Erase Cycles}]}{\text{Write Amplification Factor}}$$

$$\text{Product Lifetime (days)} = \frac{\text{Product TBW}}{\text{Expected Daily Write (TB)}}$$

Figure 6. Terabytes Written and Product Lifetime formulas.

Proactive Monitoring for Field Issues Using Storage Diagnostics

Another workload analysis approach may be done based on statistics and diagnostic information read directly from the storage device. Examples for such diagnostic information are lifetime statistics and end-of-life notifications which can be read from storage devices in easy and standard methods in eMMC, UFS and NVMe™.

Monitoring the progress of such statistics from storage devices in the field is an extremely powerful ability that enables proactive identification of field issues in case of unexpected excessive use of the storage, defects or any other abnormalities in the system. Early identification of field issues allows the implementation of preventive actions before devices may reach their end of life and may suffer failures which can require product recall.

On top of monitoring field usage, such device diagnostics may also be used in the lab as an extra analysis tool. During workload analysis, these diagnostics can be used to fine-tune tests and qualification, identify setup issues, and analyze product behavior.

Monitoring field diagnostics is a complementary effort to the workload experiment flow. While each approach has its pros and cons as shown in Figure 7, it is recommended to utilize both techniques and apply learnings between the two.

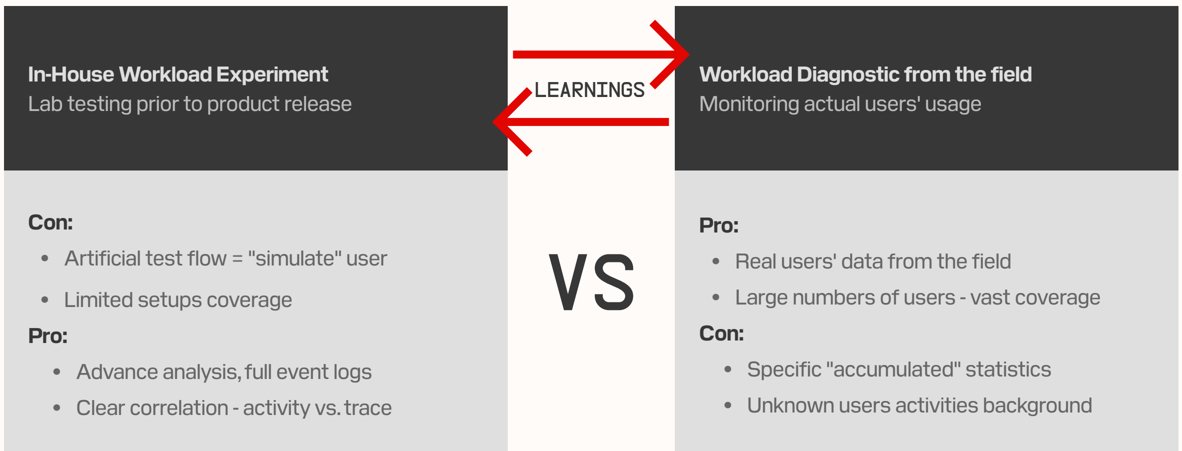

Figure 7. Comparison between lab experiments and field diagnostics approaches for workload analysis

Workload Analysis in the Automotive Market – Summary

In the continuously transforming automotive market, storage is expected to support more and more demanding applications with mixed usages than ever before. As applications change and workloads become more intense, the storage becomes increasingly critical to the reliability, durability, and user experience of the entire system.

Performing workload analysis in the lab, with Sandisk's state-of-the-art diagnostics monitoring and tools, is an imperative phase of the system validation pathway to identify unexpected behaviors, apply preventive actions and assure the storage specifications' compatibility to meet the system requirements and use cases.

Sandisk has invested greatly in tools and methodologies that enable a deep and thorough workload analysis. Using these tools, Sandisk can help OEMs and Tier-1 companies to reach higher predictability for the expected storage behavior and lifetime, and to deliver a much better solution to their end users.

For all inquiries, please email:
OEMProducts@sandisk.com
www.sandisk.com