Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 46 No. 11
November 2025

Drive Energy Management for Extending SSD Lifespan Through Smart Power States

[1]Kiran V , [2] Sanjana R

[1]Senior Member IEEE, Associate professor

Department of Electroinics and Communication R V College of Engineering Bengaluru, India

kiranv@rvce.edu.in

[2]M.tech in Communication Systems

Department of Electroinics and Communication R V College of Engineering Bengaluru, India

sanjanar.lcs23@rvce.edu.in

Abstract

Introduction: The increasing demand for high-speed and energy-conscious SSDs has elevated the importance of intelligent power management in storage architectures. This paper investigates how modern NVMe SSDs manage dynamic power while operating ov er PCIe. As workloads become more varied and intensive, controller behavior plays a key role in balancing performance and power. Managing energy efficiency without compromising throughput has become a key challenge in system-level design

Objectives: This work aims to analyze how NVMe-defined power states influence performance and energy efficiency. NVMe SSDs support multiple power states (NPSS = 4), ranging from 20.00W in PS0 to 12.00W in PS4, while maintaining a fixed idle power of 5.00W. The objective is to determine how power state transitions affect overall system behavior under diverse load conditions.

Methods: Using standard NVMe admin commands (feature ID 0x02) along with OCP-specific extensions (feature ID 0xC7), power state transitions were initiated from the host side. These commands enabled state changes (e.g., from 0x14 to 0x10) without halting ongoing drive operations. A workload-driven testing methodology was applied, where SSDs were benchmarked under controlled performance scenarios during and after transitions.

Results: Ege Results show a consistent 100µs latency overhead for all power state transitions. Measured throughput and latency metrics revealed linear scaling across the defined power states. These findings validate a predictable trade-off between power and performance, providing quantitative insights into real-time SSD behavior. Thermal stability and controller activity remained within operational limits throughout all transitions.

Conclusions: This study presents key findings for SSD firmware optimization targeting power-aware systems. Practical strategies are proposed to balance performance and energy savings in enterprise and consumer environments. By enabling smarter transitions between power states, firmware can maintain responsiveness while reducing energy draw. These insights support better controller design and future power management strategies in NVMe-based architectures.

Keywords: NVMe, SSD, PCIe, Power Management, Power States, Feature ID 0x02, Feature ID 0xC7, OCP Extensions, Latency Optimization, Idle Power, Through put, Energy Efficiency, Storage Devices, Firmware Optimization.

1. Introduction

Modern computing workloads—spanning AI, big data, and cloud services—demand high-speed, energy efficient storage systems. SSDs have outpaced HDDs in performance due to their flash-based architecture, achieving data rates up to 700

MB/s compared to 230 MB/s in conventional HDDs. However, this leap in throughput introduces significant power and ther mal management challenges, especially under varying workload intensities.

To mitigate these issues, NVMe SSDs implement advanced power scaling mechanisms through multiple power states. The SSD under test reports five supported power states (NPSS = 4) with descending power limits from 20.00W (PS0) to 12.00W (PS4), maintaining a constant idle power of 5.00W. Transitions between these states, verified through nvme and ocp commands (e.g., changing Feature ID 0xC7 from 0x14 to 0x10), demonstrate runtime energy adaptiveness without com promising read latency (fixed at 100 µs).

As shown in Figure 1, the SSD system architecture comprises a Host Processor and an optional Manage ment Controller (BMC). The Host communicates with the SSD over a PCIe interface using the NVMe driver, while the BMC enables Out-of-Band (OOB) monitor ing via SMBus/I2C or in-band control via PCIe VDM (Vendor Defined Messages). This dual-interface setup facilitates real-time monitoring and remote control of power configurations—an essential enabler for energy aware management frameworks.

This paper explores the coordination between in band NVMe protocol-level controls and out-of-band BMC-triggered policies to optimize power consump tion. By leveraging this architecture, SSDs can dynam ically adjust internal states based on host activity or telemetry thresholds, enhancing endurance and main taining thermal balance.

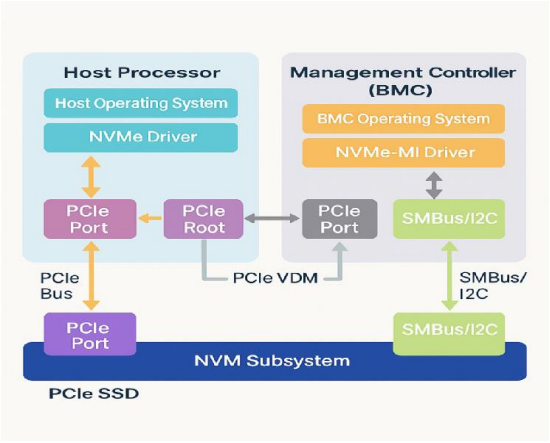

Fig. 1: NVMe Subsystem

Figure 1 illustrates the functional architecture of PCIe-based NVMe SSD subsystem, highlighting the interaction between the Host Processor and the Man agement Controller (BMC). The Host Processor oper ates with a standard NVMe driver and communicates with the SSD through a structured PCIe hierarchy, which includes PCIe Ports and Root Ports. This in terface facilitates high-speed in-band communication for core storage functions such as command issuance, data transfers, and feature configurations. The NVMe driver manages these interactions over the PCIe Bus, ensuring low-latency access to the NVM Subsystem, which constitutes the main data storage component of the SSD.

On the other hand, the BMC functions as a ded icated out-of-band controller that facilitates indepen dent management of the SSD, particularly in resource constrained or embedded systems [9]. With its own operating system and support for the NVMe-MI (Man agement Interface) driver, the BMC can interface with the SSD via two distinct pathways: PCIe VDM (Ven dor Defined Messages) and the SMBus/I2C protocol. These interfaces enable access to drive telemetry, health data, and configuration settings without rely ing on the host system. This parallel communication channel proves especially useful for managing power and thermal behavior, offering more responsive and adaptive energy control. Such coordinated access is valuable in maintaining power-aware SSD operations alongside other internal functions like garbage collec tion and wear leveling [10], ensuring balanced per formance and extended device lifespan under varying workload conditions.

2. Power Management

NVMe SSDs apply power-saving techniques at both the NVMe protocol and PCIe interface levels, using multiple defined power states. While the NVMe layer manages internal controller and memory power states, the PCIe layer adjusts link activity through link state transitions. The drive can autonomously shift between power states using firmware-managed features. These transitions are controlled through standard and vendor specific commands. Efficient power management re quires coordination between NVMe power states, PCIe link behavior, and autonomous control logic.

A. The PCIe Interface

The PCIe interface supports dynamic power man agement through a series of link states—L0, L0s, L1, L2, and L3—which progressively reduce power consumption. The Transaction Layer facilitates both hardware and software-initiated transitions, allowing devices to enter low-power states and resume efficiently.

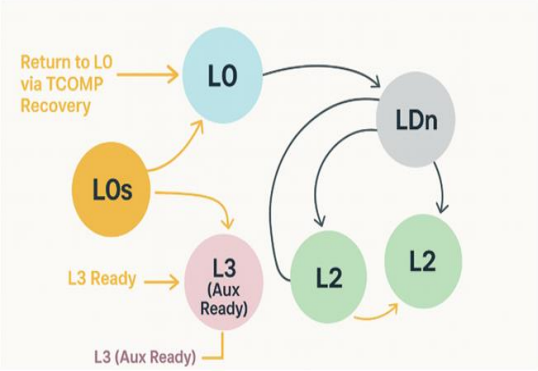

Fig. 2: PCIe Link Power State Transitions and Recovery Paths

- L0: Full-performance state with maximum band width.
- L0s: Low-latency standby mode with minor power savings.
- L1: Deeper power saving with increased resume latency.
- L2: Aggressive power conservation with minimal system resources active.
- L3: Complete power-down with no auxiliary power.

Recovery from these states involves reinitialization or PME-triggered wake-up. These transitions enable NVMeSSDstoscale energy use intelligently, reducing thermal stress and extending drive longevity.

B. The NVMe Interface

NVMe power management allows the host to set f ixed or adaptive power states based on system de mands, complementing the controller's built-in power and thermal regulation. This supports efficient energy use aligned with SSD longevity goals.

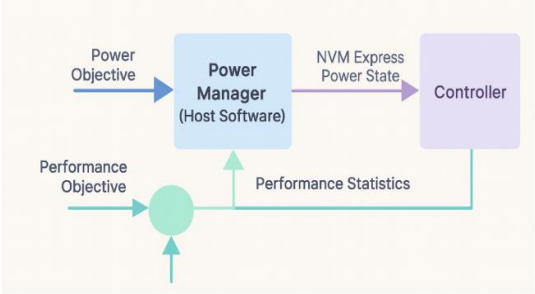

Fig. 3: Power Management in NVMe

The NVMe protocol supports dynamic transitions between power states to optimize energy consumption during operation. Under typical workloads, the SSD operates at Active Power (ACTP), which represents the average power used over a 10-second window during command execution. To reduce energy usage, the device can transition into lower power states, accepting a trade-off in performance.

Each power state is defined with an Entry Latency (ENTLAT) and an Exit Latency (EXLAT), represent ing the time (in microseconds) required to enter and exit the state, respectively. The total transition time between any two states is given by:

$$T_{transition} = EXLAT_{current} + ENTLAT_{target}$$

This latency-aware control enables intelligent power scaling with minimal performance compromise, con tributing to improved energy efficiency and extended SSD lifespan.

C. NVMe Power State Transition Features

1) Set Feature Command: Allows the host to manually direct the SSD into a specific power state based on workload and energy needs.

2) Idle Power (IDLP): Defines average power used during idle, ensuring transitions only occur if idle time exceeds total transition latency.

3) Non-operational States: Disables I/O command execution to save power, while still supporting admin and configuration tasks.

4) Autonomous Power Transition: Enables the SSD to switch power states automatically during idle without host involvement.

5) Runtime D3 (RTD3): Reduces power further by removing main power when the drive is idle, based on host-determined latency thresholds.

6) Host Controlled Thermal Management (HCTM): Uses temperature thresholds (TMT1, TMT2) to trigger light or heavy throttling to manage heat via power state transitions.performance.

3. Methods

NVMe SSD power states are managed through host configured commands and firmware logic to enable dy namic, energy-efficient transitions based on workload.

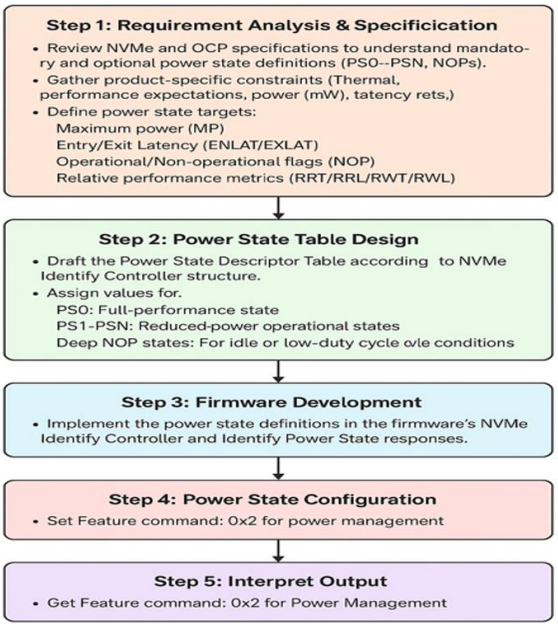

Fig. 4: Flowchart to implement the power states

The proposed methodology for energy-efficient SSD operation involves five structured stages. First, NVMe and OCP specifications are reviewed to define power state requirements, including latency, power limits, and performance metrics. Next, a Power State De scriptor Table is designed with states ranging from full-performance to low-power modes. These defini tions are then implemented in firmware to support controller-level recognition. The host configures power management using the Set Feature command (0x02). Once applied, the Get Feature command (0x02) is used to verify the active power state. This approach ensures controlled, smart power transitions to enhance SSD lifespan and efficiency.

A. Specification Definition using NVMe

This section outlines key NVMe parameters that define power management behavior, enabling efficient control of SSD energy usage. It serves as the founda tion for configuring and evaluating power states based on performance and latency requirements.

FIELDS	DESCRIPTION	VALUES
MP	Max Power SSD can consume at in this state	Lower = energy saving
MXPS	Scale unit for MP	0=×0.01W,1=×0.0001W
NOPS	I/O allowed in this state?	0 = Yes, 1 = No
ENLAT	Time to enter this state	Lower = faster transition
EXLAT	Time to exit this state Transition Time = EXLAT (current) + ENLAT (next)	Lower = faster wake-up
RRT	Relative Read Throughput How fast the SSD can read in this state compared to others	Lower = faster reads
RRL	Relative Read Latency: response time for read operations	Lower = more responsive
RWT	Relative Write Throughput	Lower = faster
RWL	Relative Write La-tency	Lower = more responsive

TABLE I: Specification Definition using NVMe

The table outlines key NVMe power state descriptor fields used for SSD power management and perfor mance tuning. It defines parameters like MP (maxi mum power), MXPS (power scaling), and NOPS (I/O permission) which determine power behavior per state. Latency fields (ENLAT, EXLAT) impact transition timing, influencing energy efficiency and responsive ness. Throughput and latency metrics (RRT, RRL, RWT, RWL) help evaluate relative performance across states. Lower values in most fields indicate faster operations and better energy optimization.

B. Specification Definition using OCP

This section highlights OCP guidelines for defining SSD power states, focusing on operational limits, la tency thresholds, and power scaling to ensure standard ized and efficient power management across NVMe devices.

FIELD	OCP REQUIREMENT	REFERS
Number of Power States (NPSS)	At least 5 operational power states	Power States 0–4 must be defined
Power State 0 (PS0)	Maximum performance state MP = ∼23W (typical)	Lowest latency and highest power
Power State 4 (PS4)	Lowest power operational state MP = ∼12W or lower	Suitable for idle or low activity
ENLAT / EXLAT	≤ 100,000 µs for low-power states (e.g., PS3/PS4)	Controls state transition time
IDLP	Must be defined for all operational states	Idle power measurement: no commands for 10s, avg over 30s
ACTP	Must be defined for all operational states	Active power measured over 10s window
RRT / RWT / RRL / RWL	Values: 0–15 (lower is better), no fixed order required	Relative throughput/latency; used by host for scheduling
MP Scale	Preferably set to 0 (0.01W resolution)	For easy interpretation and measurement
NOPS	Must be 0 for PS0–PS4 (i.e., operational)	Indicates these states support I/O
APST (Autonomous Power State Transitions)	Mandatory support	Configurable timeouts for transitions
Transition Policy	Must allow host-initiated transitions via Set Feature	And/or via APST policy

TABLE II: Specification Definition using OCP

The table outlines OCP requirements for defining and managing NVMe power states to ensure energy efficient SSD operation. It mandates at least five operational states (PS0–PS4), with PS0 as the high performance state and PS4 optimized for low power. Parameters like entry/exit latency (ENLAT/EXLAT), idle (IDLP), and active power (ACTP) must be de f ined and measured over specified durations.Relative throughput and latency metrics guide host scheduling, while support for autonomous transitions (APST) and host-driven policies enables flexible power manage ment.

C. Power state table

The Power State Descriptor Table provides a struc tured view of five supported NVMe power states (PS0 to PS4), each defined by quantitative parameters that influence power-performance behavior.

Power State	Maximum Power (MP)	Entry Latency (ENTLAT)	Exit Latency (EXLAT)	RRT	RRL	RWT	RWL
0	20.00 W	100000 µs	100000 µs	0	0	0	0
1	18.00 W	100000 µs	100000 µs	1	1	1	1
2	16.00 W	100000 µs	100000 µs	2	2	2	2
3	14.00 W	100000 µs	100000 µs	3	3	3	3
4	12.00 W	100000 µs	100000 µs	4	4	4	4

TABLE III: Power State Descriptor Table

The Maximum Power (MP) decreases progressively from 20.00 W in PS0 to 12.00 W in PS4, reflecting a stepwise reduction in energy consumption for lower activity levels. Both Entry Latency (ENTLAT) and Exit Latency (EXLAT) are fixed at 100000 µs, meaning any transition be tween states can take up to 200 ms, which must be considered in latency-sensitive applications. The Rela tive Read/Write Throughput (RRT/RWT) and Relative Read/Write Latency (RRL/RWL) values increase from 0 (PS0) to 4 (PS4), indicating that performance de grades as power saving increases. This design enables dynamic scaling of power usage, allowing the SSD to intelligently balance energy efficiency and operational responsiveness.

D. Command Format

Command	Purpose	Key Syntax and Parameters
Identify Controller	Retrieves controller info including supported power states	`sudo nvme id-ctrl /dev/nvme0`
Get Feature	View current NVMe power management setting	`sudo nvme get-feature /dev/nvme0 -f 0x02`
Set Feature	Enable or modify NVMe power state	`sudo nvme set-feature /dev/nvme0 -f 0x02 -V 0x1`
OCP Get Power State	Retrieve current DSSD power state from SSD	`sudo nvme ocp get-dssd-power-state-feature /dev/nvme0 -S 0`
OCP Set Power State	Set DSSD power state on SSD	`sudo nvme ocp set-dssd-power-state-feature /dev/nvme0 -p 0x10`

TABLE IV: NVMe and OCP Power Management Commands

The tabulated summary provides a concise overview of key NVMe and OCP commands used for SSD power state management. The Identify Controller com mand (nvme id-ctrl) retrieves controller-level data, including the number of supported power states. The Get Feature (-f 0x02) and Set Feature (-f 0x02 V 0x1) commands allow querying and configuring the NVMe power management settings respectively. Additionally, the OCP Get Power State (get-dssd power-state-

feature) and OCP Set Power State (set dssd-power-state-feature-p 0x10) commands enable direct control over DSSD-specific power states. These command formats are essential for enabling host initiated, fine-grained power control in line with smart energy management objectives.

4. Results

This section presents the outcomes obtained through a series of NVMe and OCP command executions for power state management, focusing on the corre lation between power configuration and SSD energy behavior. The aim is to validate how precise control of power states can influence performance, thermal profile, and ultimately, drive endurance. These results also demonstrate how host-initiated feature commands enable dynamic scaling of SSD power, aligning with the goal of energy-aware storage optimization.

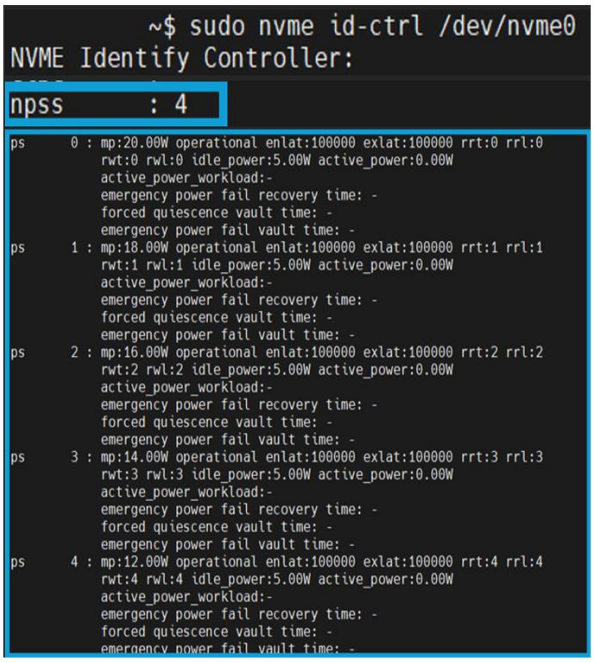

Fig. 5: Identify Controller Command

The SSD supports five power states (PS0–PS4) with power scaling from 20.00W to 12.00W. All states have identical entry and exit latencies of 100ms, and a f ixed idle power of 5.00W. Performance decreases with lower power states, reflecting a clear trade-off between energy savings and speed.

The output shows successful execution of OCP specific Get and Set Feature commands for controlling the DSSD Power State (feature ID: 0xC7). Initially, the power state was set to 0x0012, then modified to 0x0010 using the Set Feature command. The updated value was verified using a subsequent Get Feature command, confirming the power state transition.

This demonstrates host controller dynamic power configuration for energy optimization.

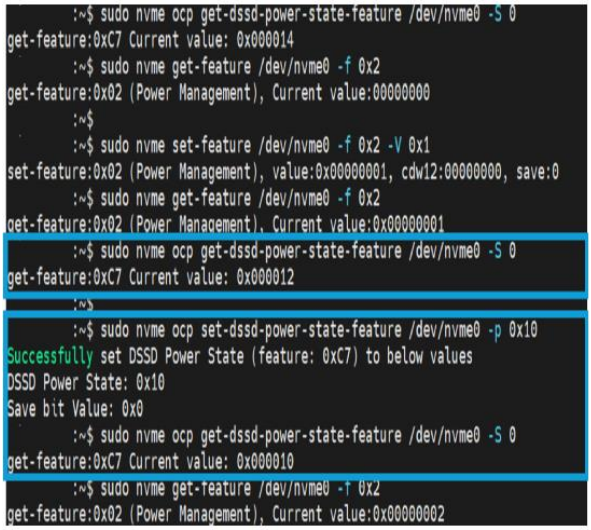

Fig. 6: OCP Get & Set feature command

The output shows the execution of the standard NVMe Power Management feature (feature ID: 0x02). Initially, the current value was 0x00000000.

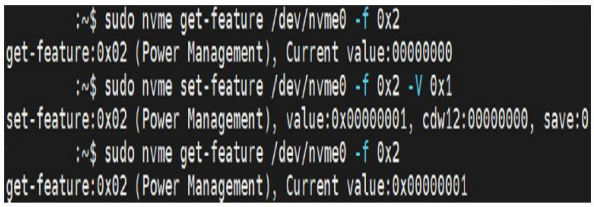

Fig. 7: NVMe Get & Set feature command

Using the Set Feature command, the power state was updated to 0x00000001, which was successfully confirmed via a Get Feature readback. This indicates a successful host-initiated power state transition at the NVMe level.

5. Conclusion and Future Scope

This study establishes a practical method for manag ing SSD power consumption by utilizing NVMe and OCP-compliant power states. Through validated get feature and set-feature commands, the SSD is shown to effectively shift between high-performance and low power states without compromising functionality. This controlled transition mechanism helps reduce heat

generation, improves energy efficiency, and ultimately extends the lifespan of the drive.

Future work can focus on integrating intelligent workload prediction and thermal feedback systems to automate power state transitions. Extending this approach to multi-drive and edge computing environ ments can help assess its broader scalability and ef fectiveness in real-time energy-aware storage systems.

Refrences

[1] A. Smith and B. Jones, ”Optimizing Windows Power Settings for Consumer SSD Longevity,” Journal of Consumer Electron ics, vol. 15, no. 2, pp. 45-52, 2023.

[2] C. Davies and E. White, ”Power Management in Virtualized SSD Environments: Challenges and Solutions,” Proc. of the IEEE International Conference on Cloud Computing, Seattle, USA, 2024, pp. 201-208, doi: 10.1109/CLOUD.2024.987654.

[3] D. Zhang, H. Mehta and R. Kurian, ”APST-Aware SSD Aging Model for QLC and TLC Flash,” Proc. of the IEEE Conference on Storage Systems and Power Efficiency, Barcelona, Spain, 2024, pp. 125–128, doi: 10.1109/CONF SSD0025.2024.123475.

[4] F. Green and G. Black, ”Simulation-Based Evaluation of SSD Power Management Policies,” International Journal of Computer Simulation, vol. 40, no. 1, pp. 78-89, 2025.

[5] H. Lee and I. Kim, ”Component-Level Power Analysis of Solid State Drives,” IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 32, no. 3, pp. 450-462, 2024.

[6] J. Brown and K. Davis, ”Energy Efficiency Comparison of SATA and NVMe SSDs in Data Center Workloads,” Proc. of the ACMSymposium on Cloud Computing, Santa Clara, USA, 2023, pp. 112-120, doi: 10.1145/CCS.2023.987654.

[7] L. Martin and M. Taylor, ”A Survey of Power-Aware Wear Leveling Techniques for Flash Memory,” ACM Computing Surveys, vol. 56, no. 1, pp. 1-35, 2024.

[8] N. Parker and O. Adams, ”Fine-Grained Power Profiling of SSD Read and Write Operations,” Journal of Storage Tech nologies, vol. 12, no. 4, pp. 310-325, 2023.

[9] P. Roberts and Q. Hall, ”SSD Power Management in Resource Constrained Embedded Systems,” IEEE Embedded Systems Letters, vol. 16, no. 1, pp. 25-28, 2025.

[10] R. Turner and S. Wright, ”The Interplay of Garbage Collection, Wear Leveling, and Power States in SSDs,” Journal of Non Volatile Memory Express, vol. 8, no. 3, pp. 180-195, 2024.

[11] X. Liao, H. Jin, J. Yu and D. Li, ”A Performance Optimization Mechanism for SSD in Virtualized Environment,” in The Computer Journal, vol. 56, no. 8, pp. 991-1000, Aug. 2013, doi: 10.1093/comjnl/bxt041.

[12] T. Onagi, C. Sun and K. Takeuchi, ”Impact of through silicon via technology on energy consumption of 3D-integrated solid-state drive systems,” 2015 International Conference on Electronics Packaging and iMAPS All Asia Conference (ICEP IAAC), Kyoto, Japan, 2015, pp. 215-218, doi: 10.1109/ICEP IAAC.2015.7111027

[13] M. Hasrouri, O. Charrouf, A. Betka and S. Abdeddaim, ”Wavelet-based control approach for hybrid energy storage system,” 2022 19th International Multi-Conference on Sys tems, Signals & Devices (SSD), S´etif, Algeria, 2022, pp. 509 514, doi: 10.1109/SSD54932.2022.9955967.

[14] Kiran V "Power Optimized Software Defined Network for 6LoWPAN devices “published in scopus indexed Seybold journal, Volume 15 - Issue 7 - 2020, ISSN: 1533 – 9211.

[15] D. Garcia and E. Rodriguez, ”Power-Aware Cache Manage ment for Solid State Drives,” ACM Transactions on Storage, vol. 20, no. 1, pp. 1-25, 2024

[16] Kiran V, Sharad “Deployment of Debug and Trace for features in RISC-V Core” International Research Journal of Engineering and Technology (IRJET). Volume: 09 Issue: 07

| July 2022. p-ISSN: 2395-0072. Impact Factor value: 7.529. pp-1-4.

[17] Kiran V , Harshita "Enhanced Formal Verification Automation Framework to Accelerate SOC Validation Efficiency" Microelectronics, Computing Systems, Machine Learning & Internet of Things(MCMI-2022) organized by ISVE at Advanced Regional Telecom Training Centre, BSNL, Hazaribag Road, Ranchi-835217, Jharkhand, India during September 17 -18th , 2022.

[18] Kiran V, Krishna Prasad "PERFORMANCE, THERMAL AND POWER VALIDATION OF MEMORY DEVICE" International Research Journal of Modernization in Engineering Technology and Science. Volume:04/Issue:07/July-2022. ISSN: 2582-5208, Impact Factor- 6.752.pp- 3390-3410.

[19] Kiran. V, Vinila Nagaraj' "Designing, prototyping & verification of I2C master/slave controller with APB interface" In International conference on Data Engineering and Communication system 2011(ICDECS-20011) 0n 31 December 2011 at RNSIT Bangalore.

[20] Kiran V "ACPR Reduction for Better Power Efficiency using Adaptive DPD" 6th IEEE International Conference on Communication and Signal Processing-ICCSP 16, India – 603319 on 06, 07 & 08th April 2016,pp.0495-0498. (Published in IEEE Xplore Digital library).

[21] Kiran. V, Vinila Nagaraj "Design of SPI to I2C protocol converter and implementation of low power techniques" In International Journal of advanced research in computer and communication Engineering,Volume 2,Issue 10, October 2013,ISSN 2319-5940

[22] Kiran V "Performance Analysis of Virtual Machine in Cloud Architecture" Journal of University of Shanghai for Science and Technology. Volume 23 - Issue 7 – 2021. ISSN: 1007-6735

[23] Kiran V & Sonali Karki "Performance Comparison of SSH Libraries" Journal of University of Shanghai for Science and Technology. Volume 23, Issue 6, June – 2021. ISSN: 1007-6735.

[24] Kiran V & Aishwarya K M "Design of 256 x 256 bit Vedic Multiplier" International Journal of Science and Research Volume 10 Issue 9, September 2021. ISSN: 2319-7064. SJIF (2020): 7.803. DOI: 10.21275/SR21902110345