How New Energy-Efficient SSD Controllers Can Dramatically Improve Data Center TCO

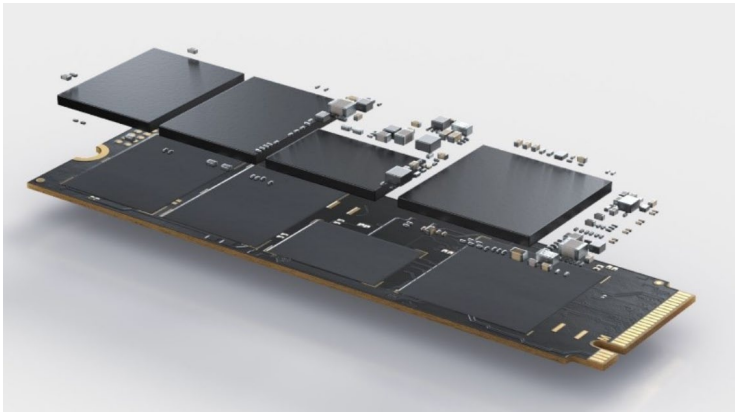
Jonmichael Hands, Sr. Director Product Planning

07/08/24

Agenda

- Power Efficiency Disparities: Data Center vs. Consumer SSDs
- Controller vs. NAND power
- Utilizing NVMe Power States
- Demo – Understanding Measured Results
- Drive Power Efficiency Impact to Server Power
- Drive Power Efficiency Impact to Rack Level TCO
- FDP Improves Write Bandwidth
- Quantifying the Impact: Sustainability vs Energy Efficiency

Power Efficiency Disparities: Data Center vs. Consumer SSDs



Data Center SSDs:

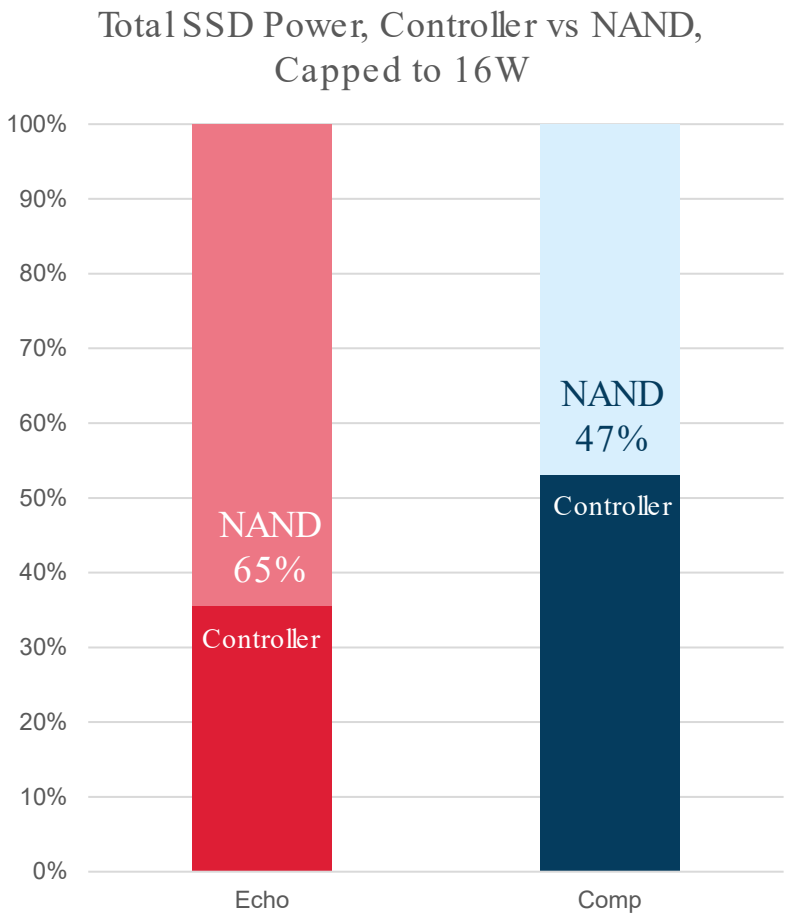
- Optimized for continuous workloads and steady-state performance
- Power can be efficiently managed by NVMe power states to cap the drive's TDP.
- The critical power efficiency metric is performance per watt in the active state and idle active (low latency resume).

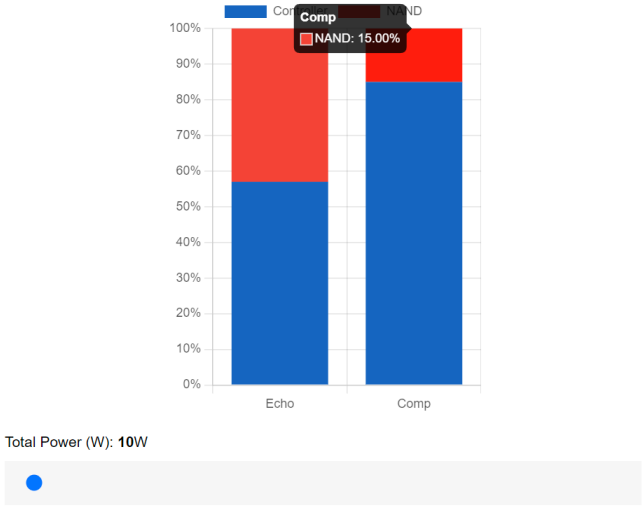
Performance per Win the active state is today's key metric for data center power efficiency.

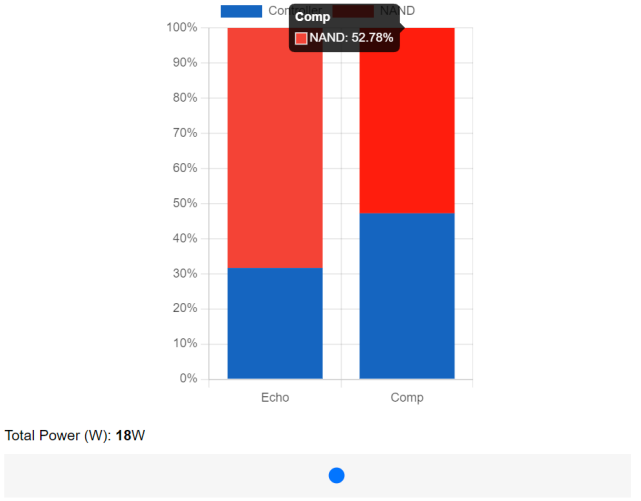
Consumer SSDs:

- typically experience bursty workloads followed by extended idle times (gaming, content creation, office productivity).
- This necessitates SSDs that can deliver responsive performance during active use while minimizing power consumption during idle periods.
- Consumer SSDs use technology like NVMe autonomous power state transitions and PCIe L1.2 to deliver near-zero idle power.

Controller vs. NAND power







The higher a controller's power consumption, the more performance loss and reduction in performance and power efficiency when the total power budget is capped

Utilizing NVMe Power States

- Discover available power states through NVMe identify controller
- Figure out what power state the drive is currently in

```
nvme get-feature /dev/nvme0 -f 2 -s 0
```

- Set a new power state that is saved and persistent across power cycles

```
nvme set-feature /dev/nvme0 -f 0x2 -v 0 -s
```

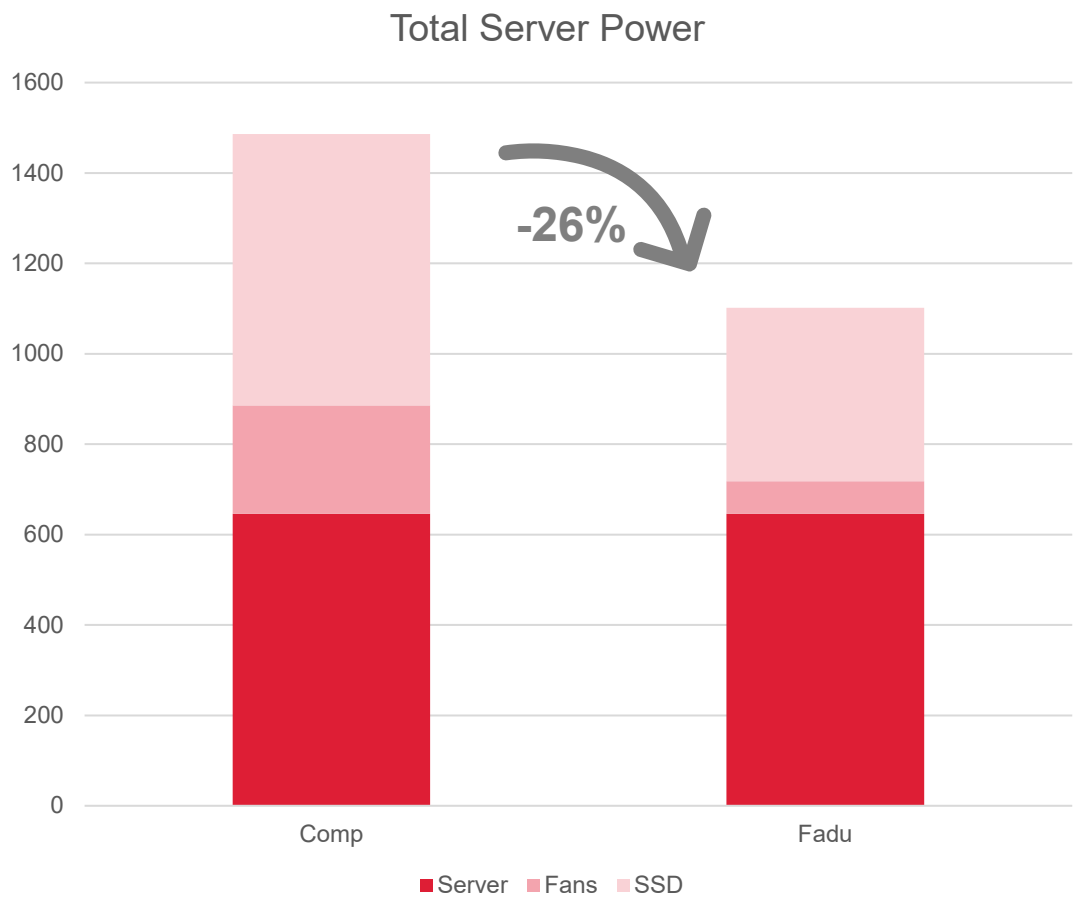

```
nvme id-ctrl /dev/nvme0n1 -H

ps      0 : mp:35.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      1 : mp:25.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      2 : mp:20.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      3 : mp:18.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      4 : mp:16.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      5 : mp:14.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      6 : mp:12.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      7 : mp:10.00W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
ps      8 : mp:8.25W operational enlat:100000 exlat:100000 rrt:0 rrl:0
          rwt:0 rwl:0 idle_power:- active_power:-
          active_power_workload:-
```

Demo – Understanding Measured Results

https://blogs.fadu.io/energy-efficient-controller/

Drive Power Efficiency Impact to Server Power

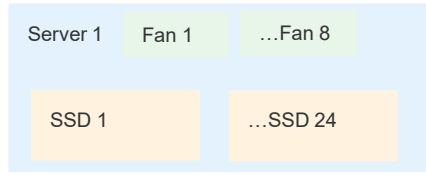


Fadu E1.S: 16W

Comp PCIe 5.0 E1.S: 25W

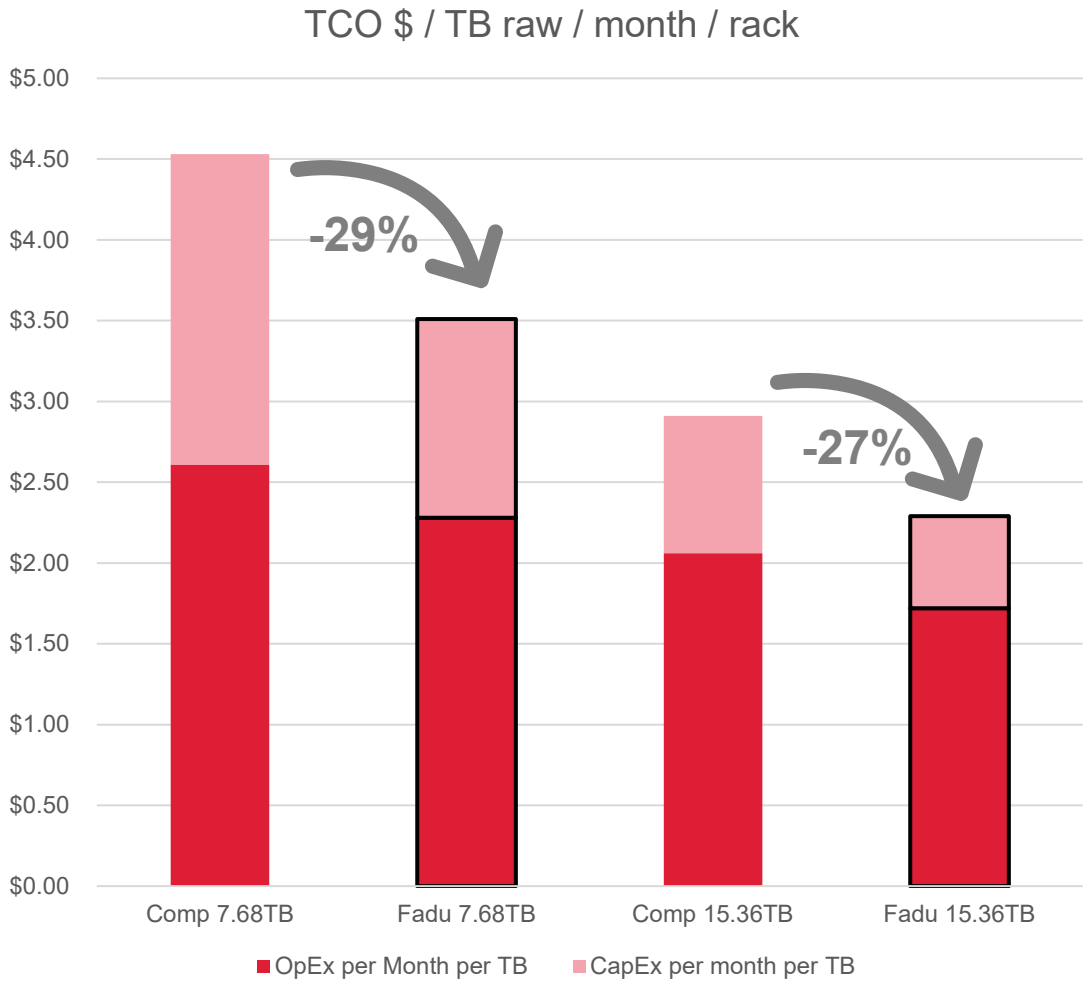
Server (CPU, DRAM, motherboard, NIC)

40mm fan: 30W



Fans consume a lot of power in servers! Small component changes make a big difference

Drive Power Efficiency Impact to Rack Level TCO

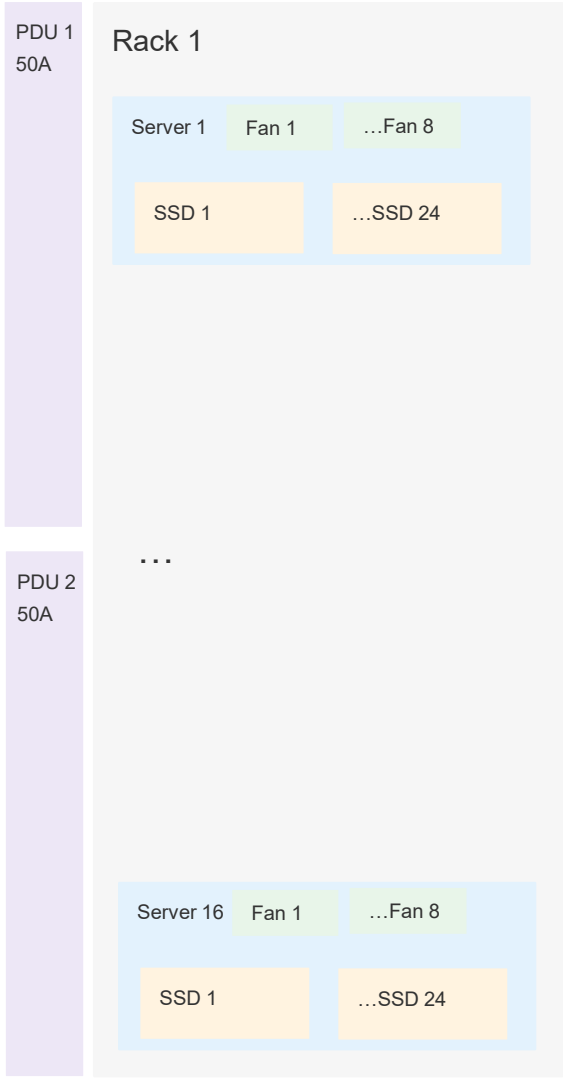


Fadu E1.S: 16W

Comp PCIe 5.0 E1.S: 25W

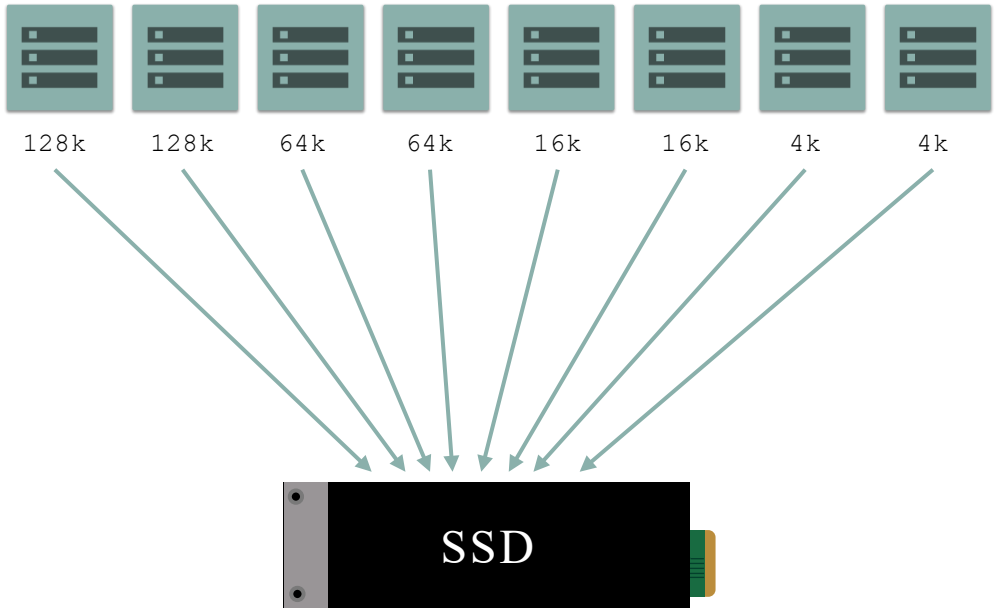
Server (CPU, DRAM, motherboard, NIC)

40mm fan: 30W

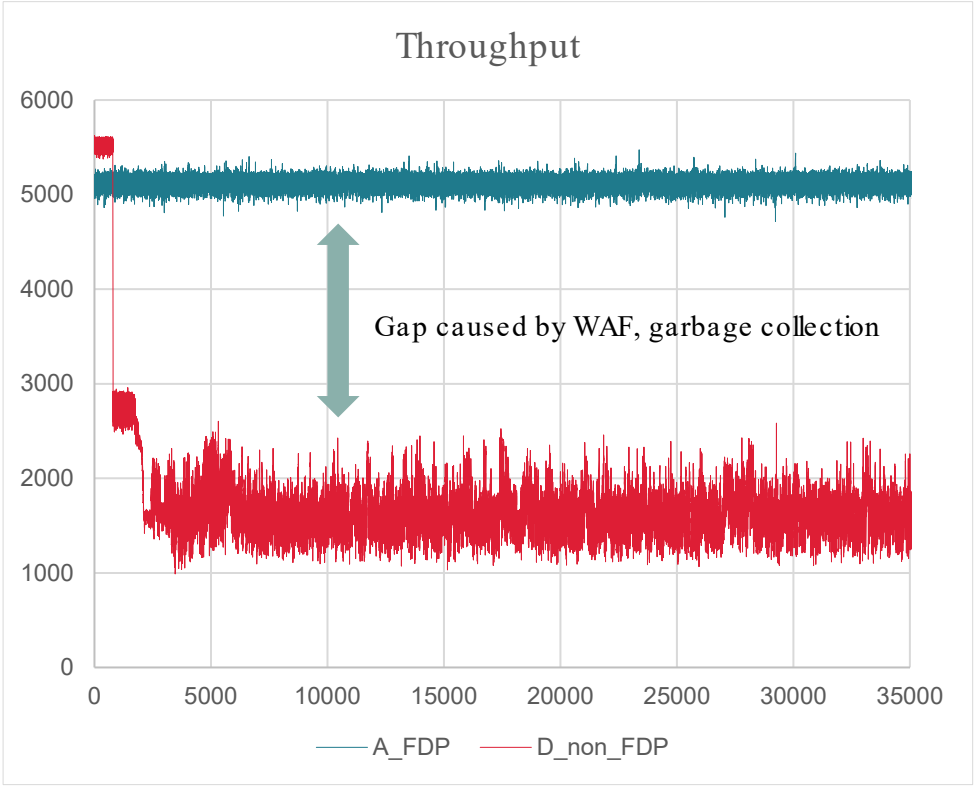


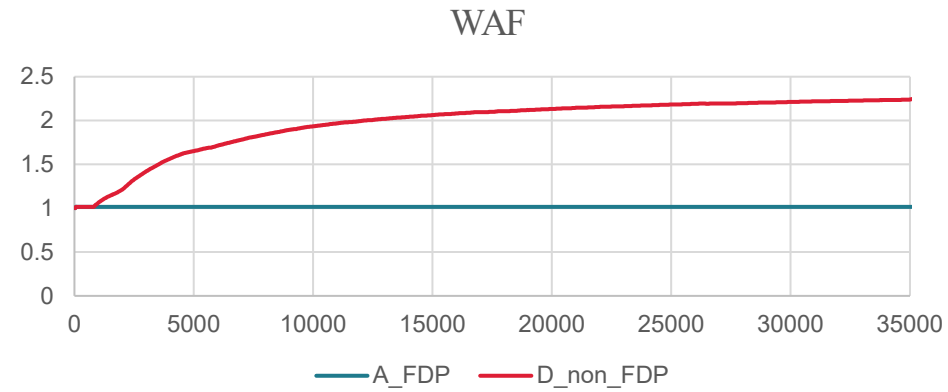
45% more servers per rack (PDU limit at ~15kW)

FDP Improves Write Bandwidth – and perf/W!



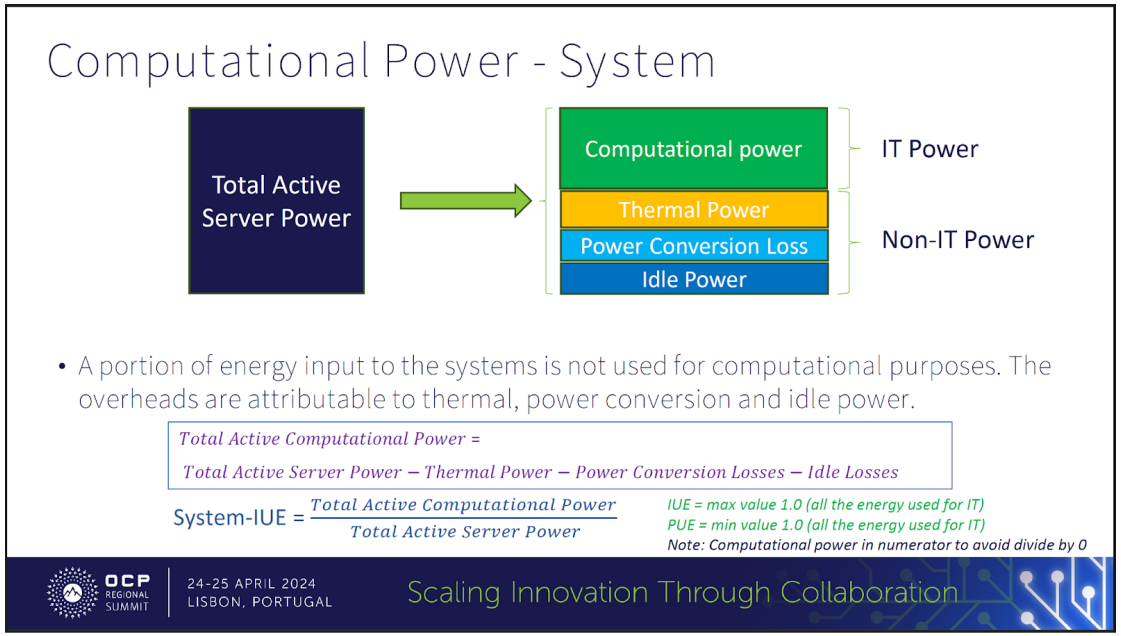
8 different write threads of different block sizes
Achieve WAF=1, full NAND bandwidth with FDP





Quantifying the Impact: Sustainability vs Energy Efficiency

- Most people confused energy efficiency and sustainability
- OCP sustainability has published whitepaper on sustainability in ICT
- New metrics, Infrastructure Utilization Efficiency (IUE) metric to classify energy consumed at system and rack level that distinguishes between compute vs non-compute power
- Life Cycle Assessments (LCAs) offer a comprehensive way to document energy efficiency



Energy efficiency is a key sustainability driver of the use phase!

Creating Future