


# Flash Usage in Cloud Datacenters

Kushagra Vaid  
GM, Server Engineering, Microsoft  
kvaid@microsoft.com

FLASH MEMORY SUMMIT, AUG 2013

# Cloud-scale services @ Microsoft

 Windows Live®

500M Active Live IDs  
*59 markets and  
36 languages*

 Windows Azure

 Microsoft  
SkyDrive.

 Hotmail.


350M Active  
Accounts

 skype

300M+ Users  
*76 markets and  
48 languages*

 bing™

Over 5.5B WW  
Queries Each  
Month

 Windows Live  
Essentials

Windows Live  
Essentials

 msn.

Over 500M  
Unique Users  
Each Month

 XBOX

25M Users

 Microsoft  
Advertising

14B Ads  
Per Month

Microsoft  
Exchange  
Hosted Services

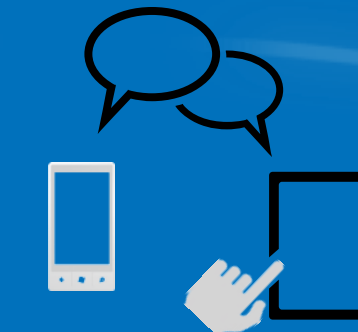
2-4 billion emails  
per day

 Office 365

200+ Cloud Services



PC/Tablets

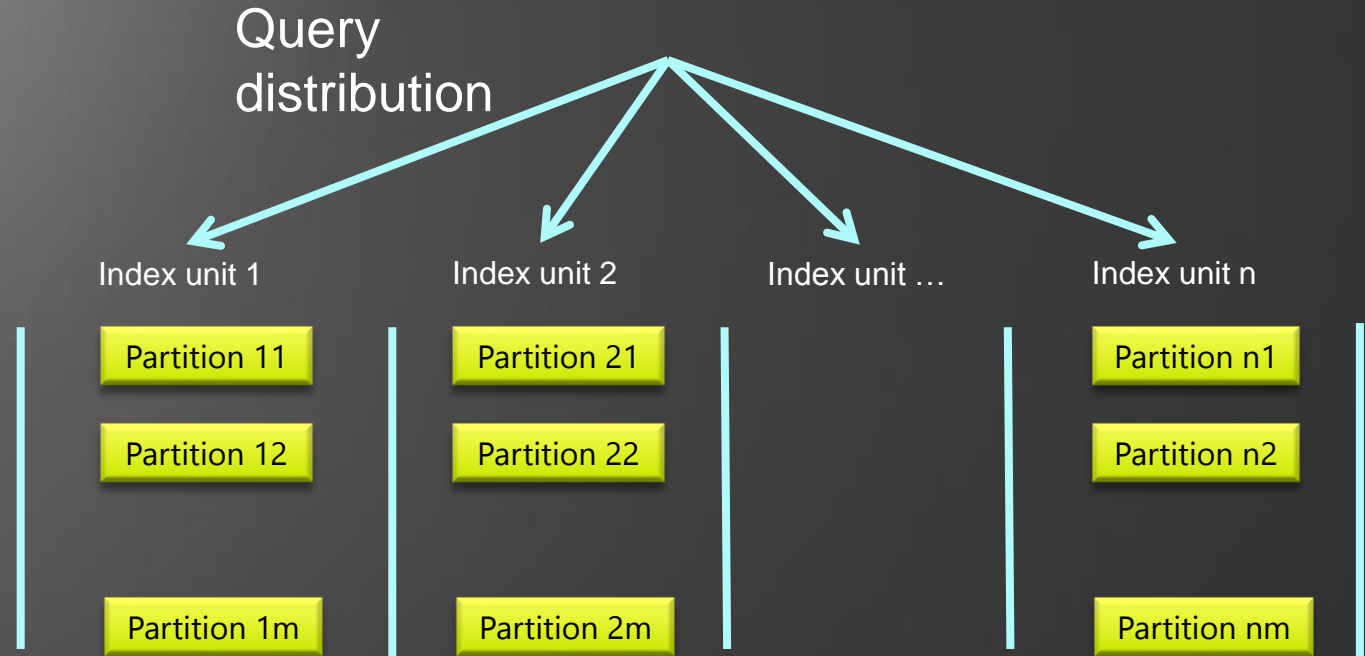
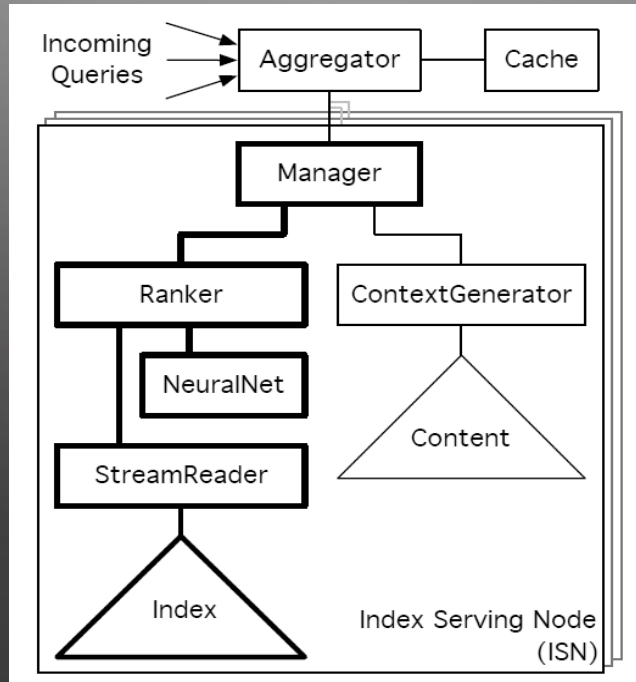


MOBILE



TV/HOME

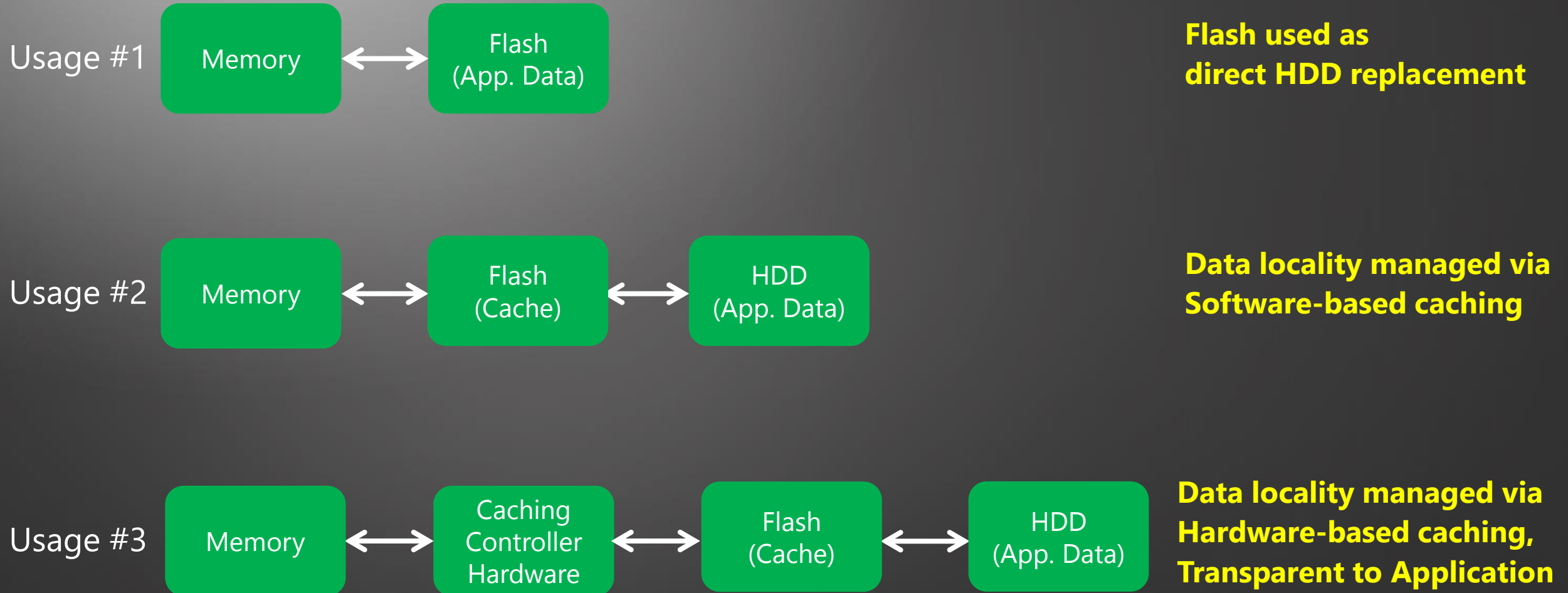
# Large scale cloud application: Web search



Source: "Web search using mobile cores, ISCA 2010"

*Query performance is measured as an aggregate of ALL query processing nodes (not just individual discrete nodes)*

# Data management (typical) using Flash devices



# Data management (typical) using Flash devices

Usage #1

***Not cost effective for large datasets***

**Flash used as  
direct HDD replacement**

Usage #2

***Most cost effective. App can be specifically tuned  
for Flash perf/endurance parameters***

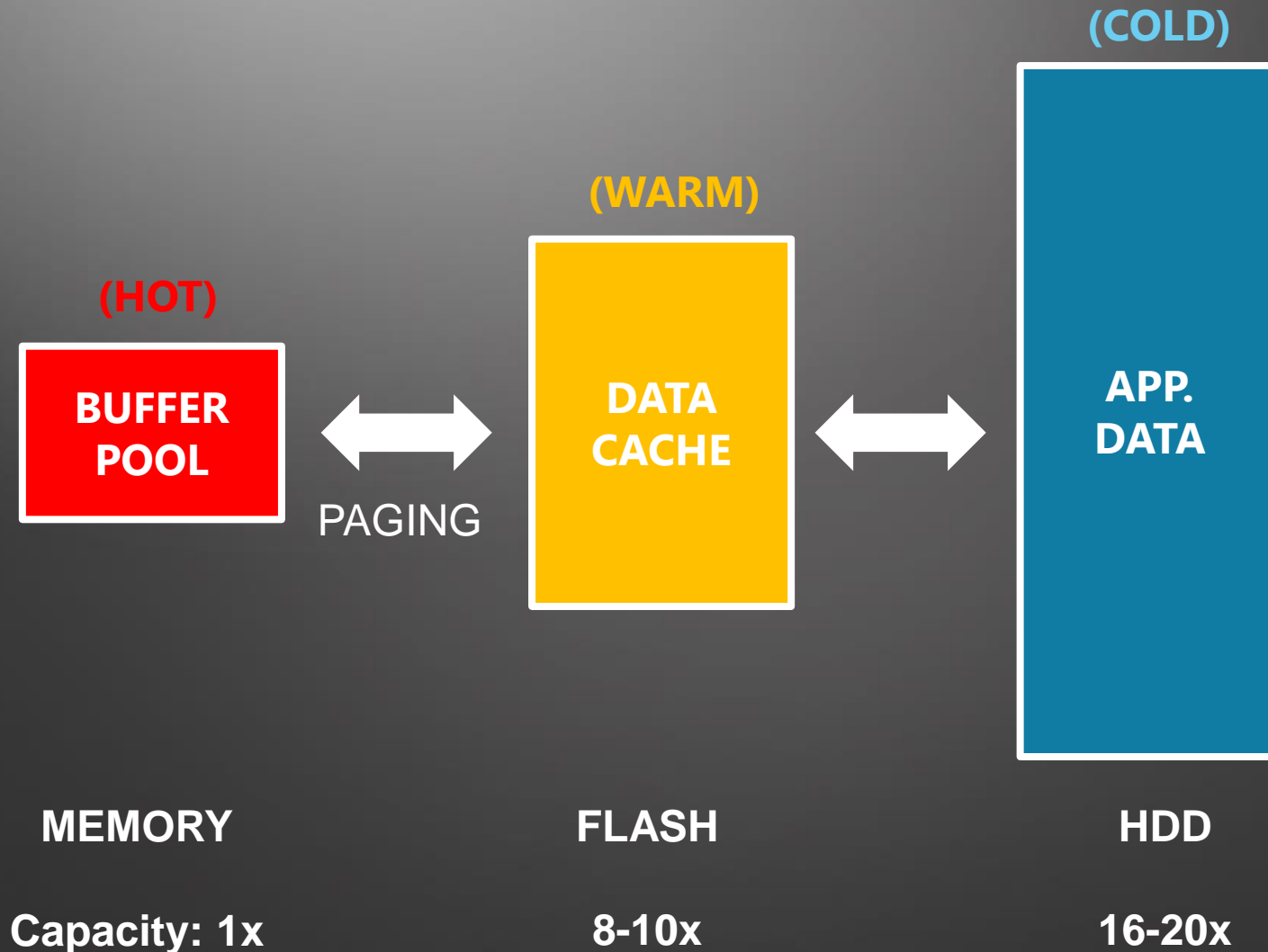
**Data locality managed via  
Software-based caching**

Usage #3

***Less cost effective than #2. Well suited for  
accelerating legacy/3<sup>rd</sup>-party Apps***

**Data locality managed via  
Hardware-based caching,  
Transparent to Application**

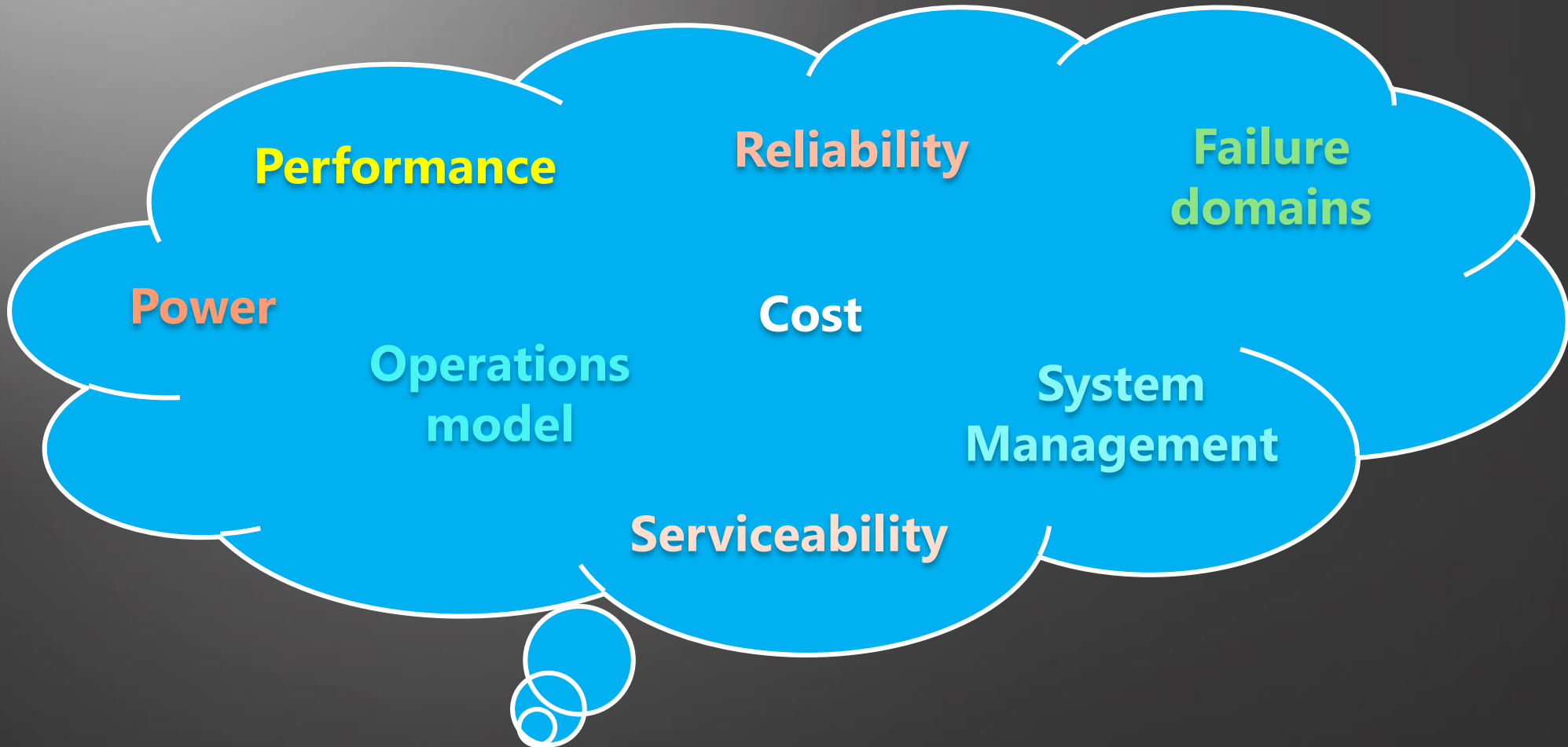
# Software-based caching details



Software is customized and tuned for ...

- Flash device latency and bandwidth characteristics
- Access patterns (random, sequential), block sizes
- Data placement and locality (hot/warm/cold migration)

# The Goldilocks problem applied to Cloud servers



*Not too "big" Not too "small" !*

# Cloud SSD desired characteristics

## Not Enterprise :

- RAS features not needed (e.g. dual ported SAS)
- MTTF = 1.5M hours (AFR ~0.5%)

## Cost economics

- Ride commodity cost curves, e.g. MLC / TLC NAND
- Connect via available motherboard ports (e.g. no expander/HBA)



# Cloud SSD desired characteristics (continued)

## Performance

- Read IOPS: ~100K peak, ~40K average is good enough
- Latency consistency: 99.9999<sup>th</sup> percentile close to average latency

Endurance : Sweet spot (15-20K P/E) and low retention (30 days max)

Low Power: ~5W max

# Challenges with current SSDs

- Endurance Metrics: Drive writes per day? Terabytes written? PE-cycles?
- Vague and inconsistent SMART counters
- Inconsistency in power-loss features
- Latency inconsistencies
- Varying performance over lifetime
- Management at scale

# Call to action

- Cloud ≠ Enterprise! Understand requirements of the Cloud
- Drive greater standardization to bring value to customers
- Performance is not the only metric that matters
- This is the right time to drive disruptive innovation !

# Q & A

