Open. Together.

OCP SUMMIT
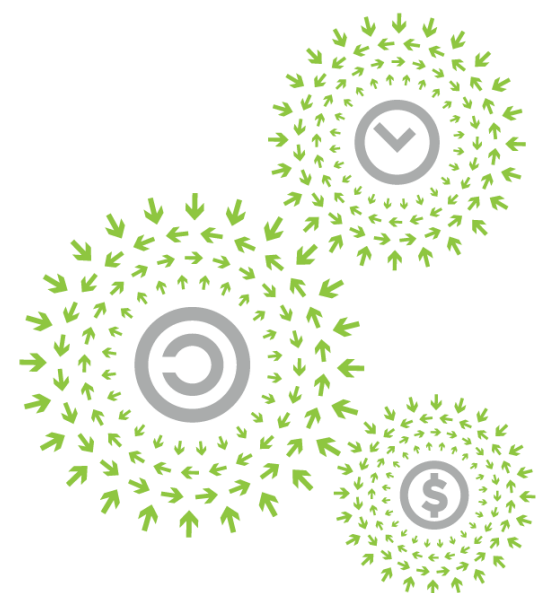
STORAGE

HDD IOPS – The Never Ending Story

Jason Adrian
Principal Program Manager
Azure Storage Hardware, Microsoft

OPEN PLATINUM™

OCP SUMMIT

Open. Together.

How many IOPS are enough?

How do you measure performance?

Throughput (Bandwidth)?

IOPS?

Latency?

IOPS vs Latency

IO performance is typically measured in IO throughput at a target latency

How many IOPS can a drive or system provide @ 30ms p99 response time?

What if that limit was 20ms or 50ms?

If you want marketing slides, you max queue depth

If you want a high performance storage platform, you want low latency!
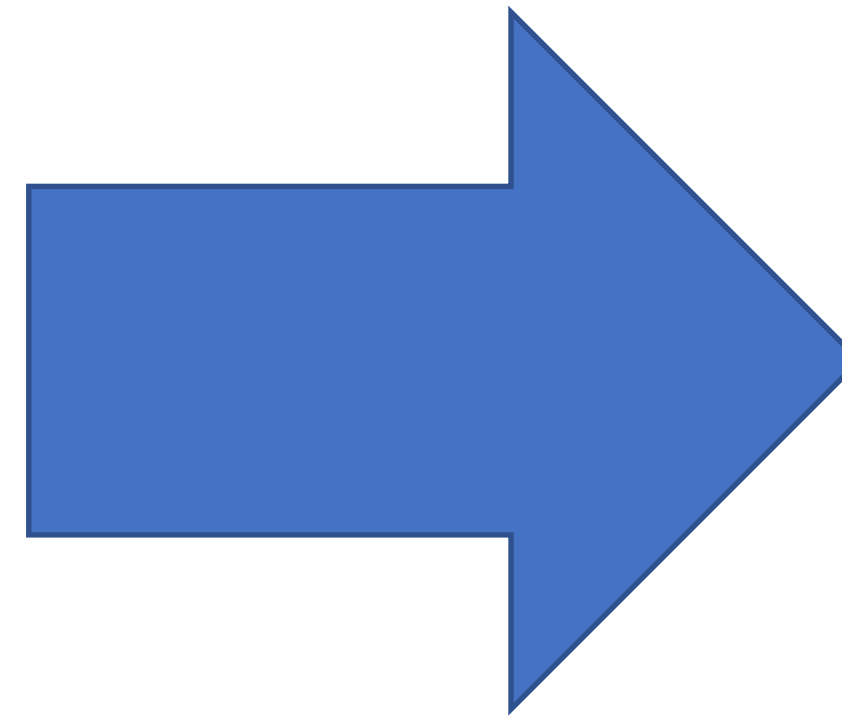
Open. Together.

But this isn't new…

Do you remember…
15k 3.5" drives?
10k 3.5" drives?
15k 2.5" drives?
10k 2.5" drives?

SSDs

Today SSDs have displaced high performance HDDs

Capacity or performance?

HDDs are still king for affordable capacity!
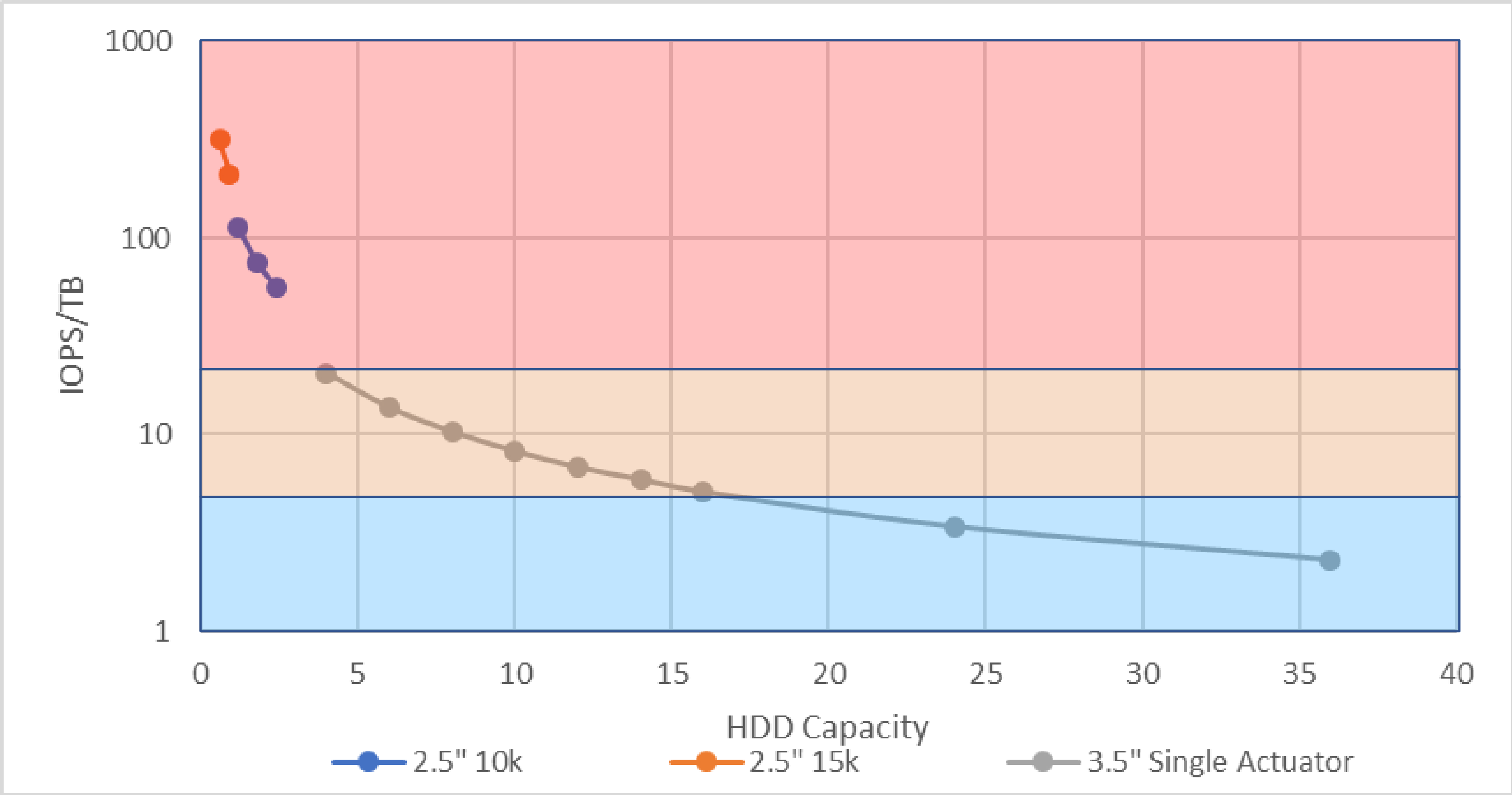
The key is using high capacity drives effectively

Do you combine warm and cool workloads?

Does your workload cool at the rate drive capacity increases?

Employ read caches in memory or SSDs

…. Or short stroke the drives (strand capacity)?

Capacity or performance?

High Performance & Low Latency Applications (SSDs)

Near-line storage Ex: Photos, analytics, ML (HDD)

Cool, Archival, Backup (HDD, Tape, Optical)

4k Random Read, QD1

Open. Together.

We're always out of IOPS...

Every new capacity point you hear "we're out of IOPS" but workarounds/optimizations are found

10TB+ drives - common pain point across the industry

- latency back to the forefront

- 16+ TB drives on the horizon, how do we adopt them?

Some have asked drive providers to bring 3.5" 10k drives back

  - Higher power & costs

Some asked for HDD firmware optimizations
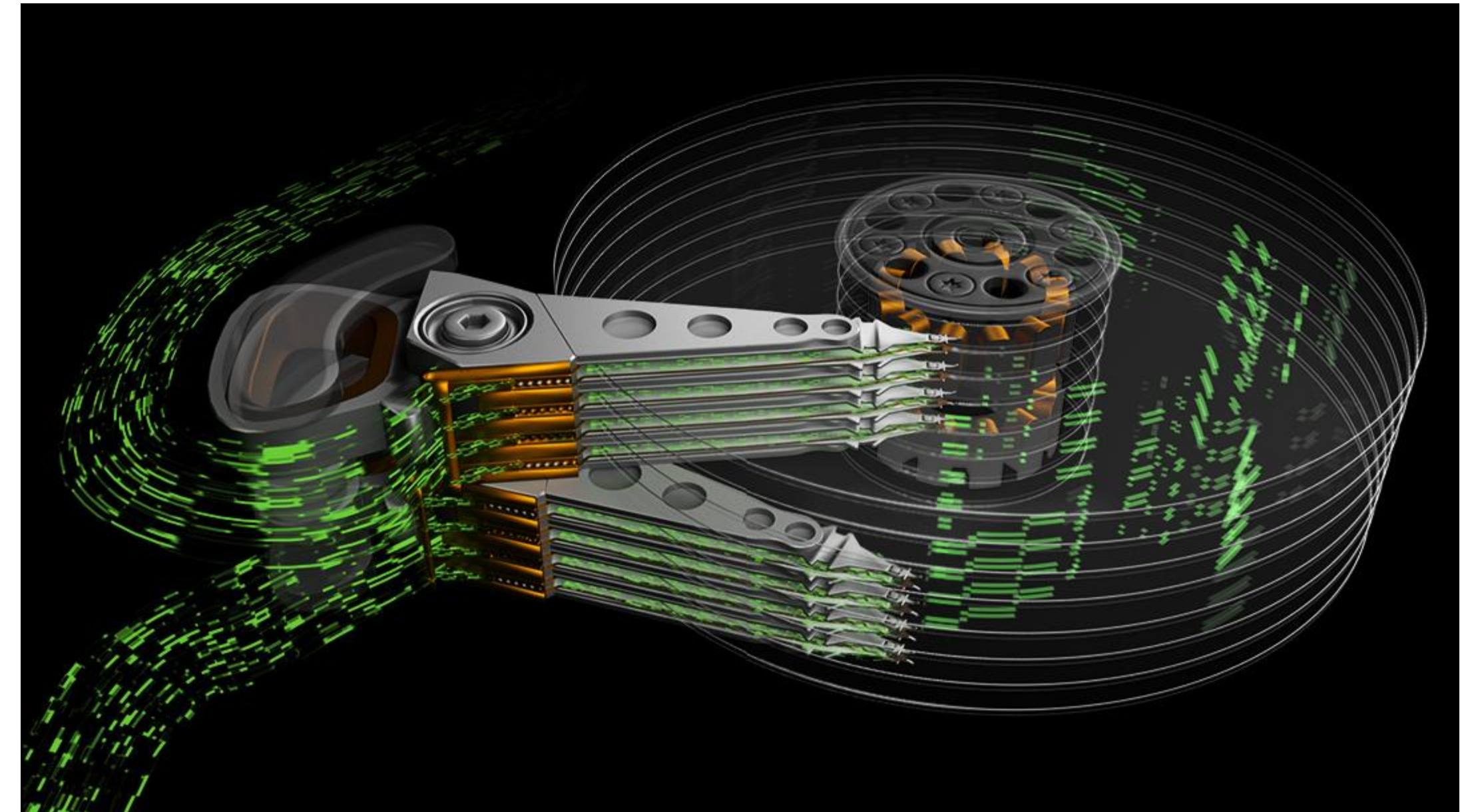
Some focused on caching

Multi Actuator Drives

Instead of increasing RPM, add more actuators!

Increase performance (IOPS and BW) nearly 2x

But… this doesn't come for free

- Higher power

- Higher cost

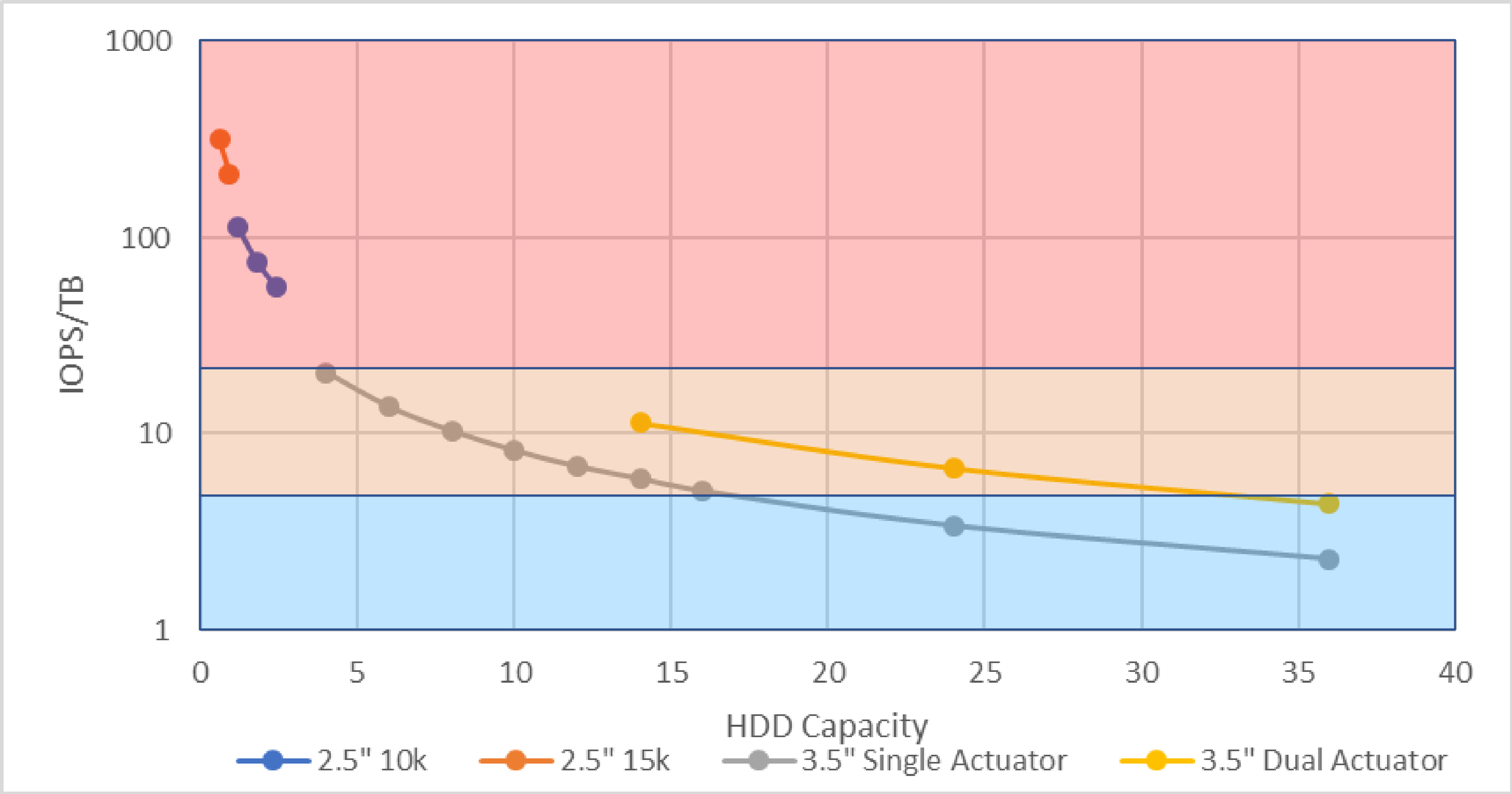- Challenging mechanical design

Seagate

Multi Actuator Drives

HDDs have promising density increases with HAMR/MAMR
With 20TB+ drives, we will need to consider all optimizations
- Queuing optimizations
- Caching optimizations
- Software optimizations
- Technologies like multi-actuator

Open. Together.

Capacity or performance?

4k Random Read, QD1

Building an Ecosystem

It's not vaporware

Seagate and Western Digital have made product/technology announcements and demo's

Seagate

Western Digital

System Readiness?

Needs chassis able to cool and power the drives
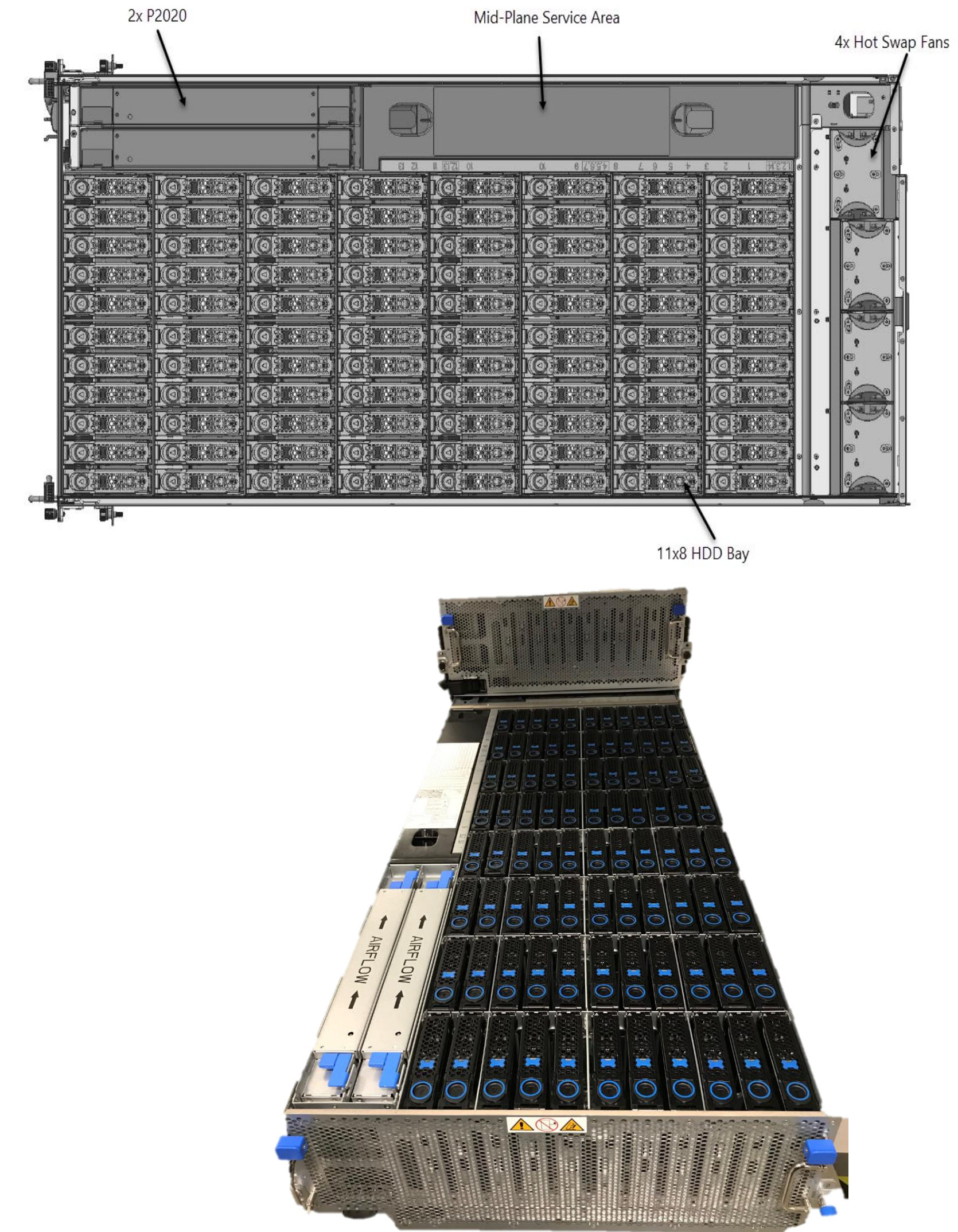
Drives are capable of high sequential streaming

- Is the storage platform and NIC sized appropriately?

Project Olympus J2010 is being validated to support these higher performance drives

Open. Together.

Olympus J2010 Support

- System was designed for high power drives (up to 13W)

- Will test against dual actuator HDDs to confirm RV/AV performance

Multi Actuator Drives

Multi actuator drives alone are not the solution!

They will decrease average latency, and increase throughput

We still need to reduce the tail latencies

- One example is Fast Fail Read

We also need to work with HDD vendors for IO Priority

- How can we share intelligence and internal queue's with the HDD schedulers?

What's next?

➤ Working together to define future HDD features and product requests

➤ For dual actuator – needs discussion on how 2 LUNS interact

➤ Propose we start a workgroup to discuss multi-actuator features and requirements

Open. Together.

OCP Global Summit | March 14–15, 2019

OCP SUMMIT