Evaluating Host Aware SMR Drives

Fenggang Wu, Ming-Chang Yang*, Ziqi Fan, Baoquan Zhang,
Xiongzi Ge, David Du
University of Minnesota, Twin Cities *National Taiwan University

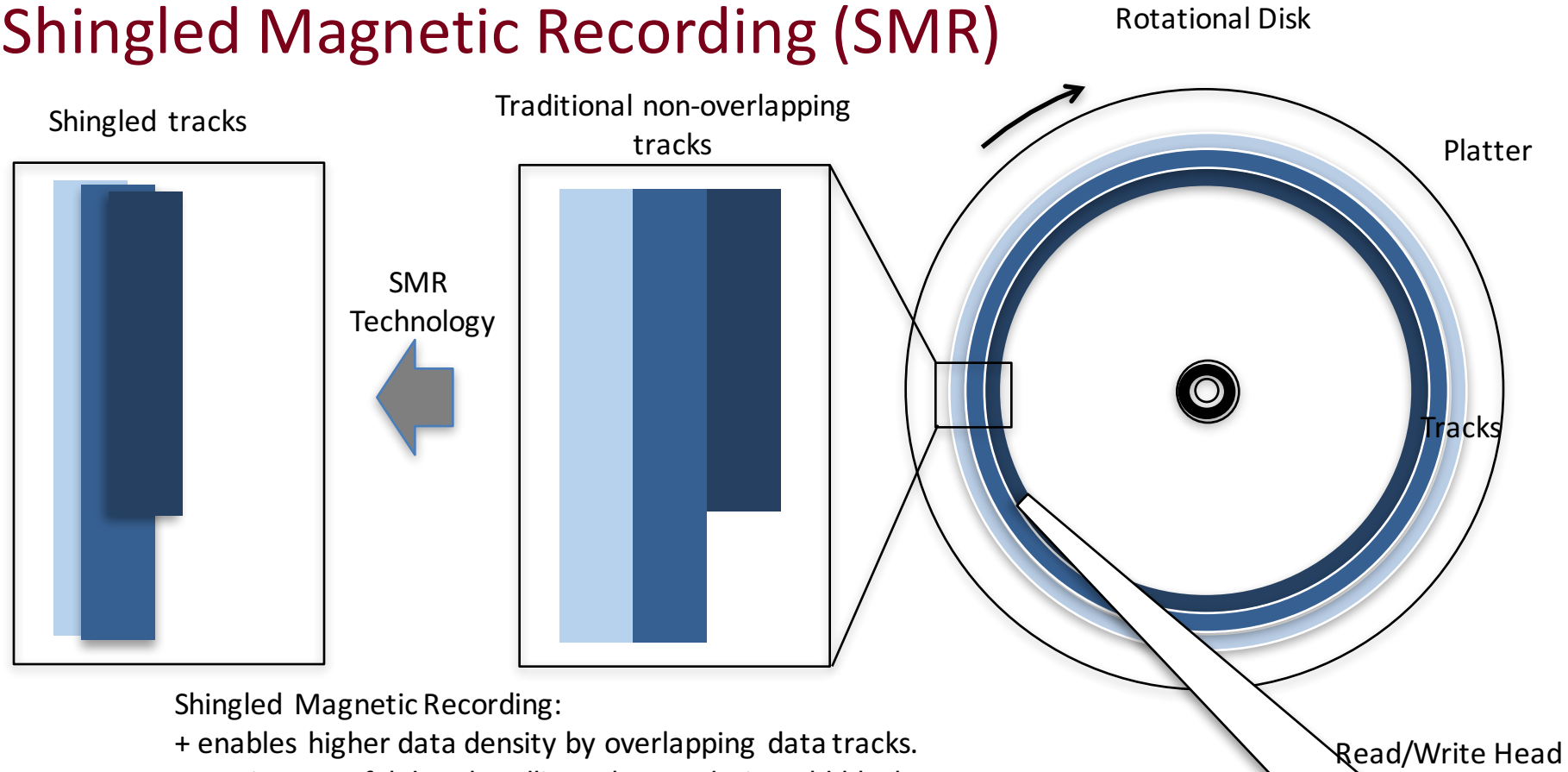
USENIX HotStorage, Jun. 20, 2016

UNIVERSITY OF MINNESOTA

Center for Research
in Intelligent Storage

National Taiwan University

Shingled Magnetic Recording (SMR)



Shingled Magnetic Recording:
+ enables higher data density by overlapping data tracks.
- requires careful data handling when updating old blocks.

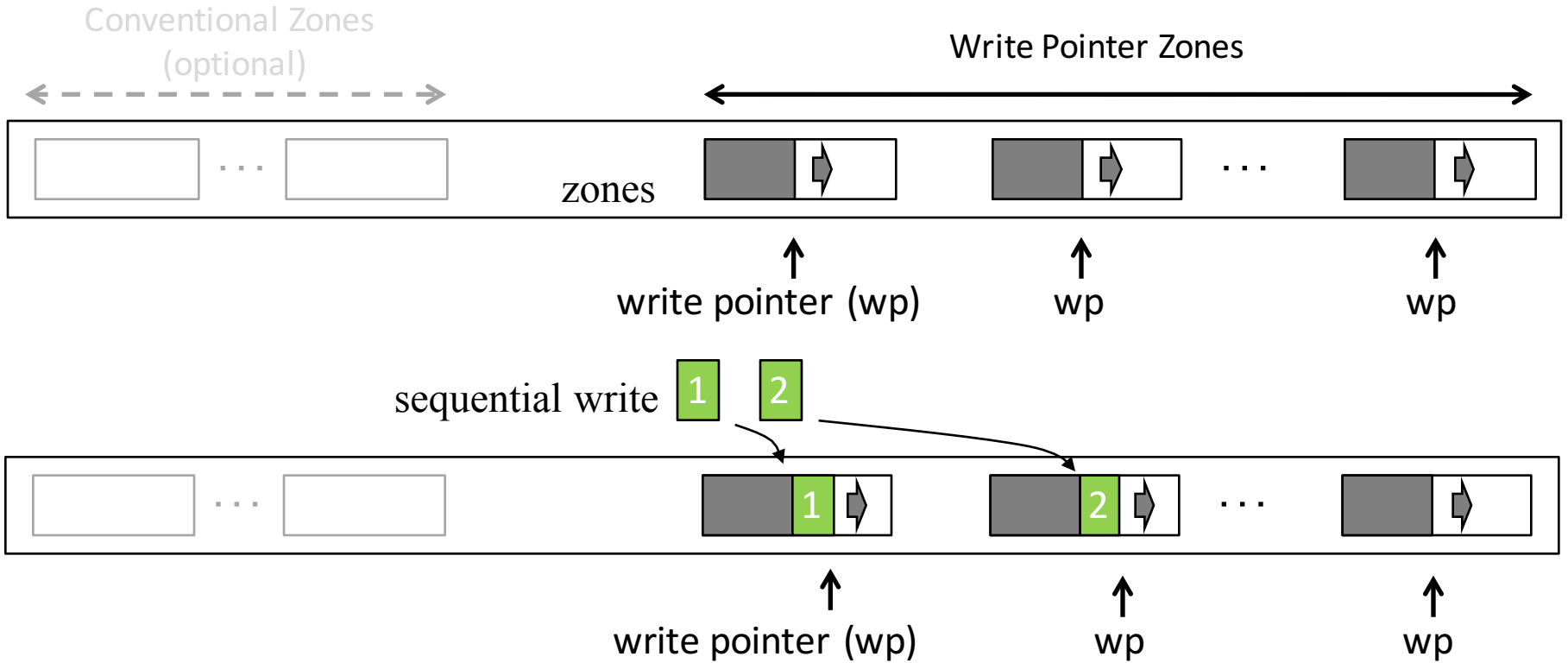
Motivation & Introduction

- Objective:
  - to build low-cost large-scale storage system using SMR drives.
- We focus on Host Aware SMR (HA-SMR) drives
  - HA-SMR is the superset of the two other SMR models (Host Managed and Drive Managed) in functionality.
- Roadmap:
  - **Understand** intrinsic features of HA-SMR drives by performance characterization.
  - **Summarize** the lesson learned and system implications from the performance testing.
  - **Explore** a Host-controlled indirection Buffer (H-Buffer) to meet the design challenge of HA-SMR systems by exploiting the special characteristics of HA-SMR.

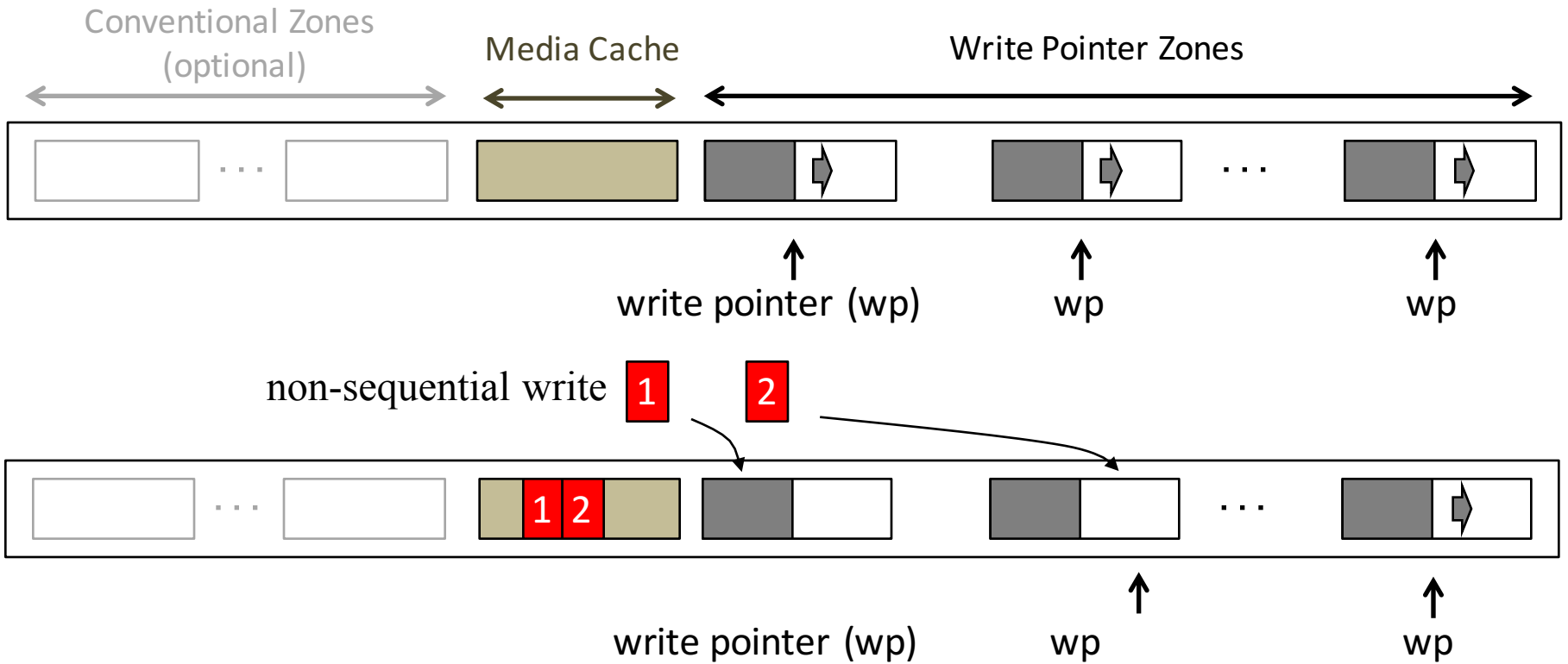
Outline

- Motivation & Introduction
- HA-SMR background
- Testing Results and Lessons Learned
- Host-controlled Indirection Buffer (H-Buffer)
- Related Work
- Summary & Future Work

HA-SMR Drive Background



HA-SMR Drive Background (Cont'd)



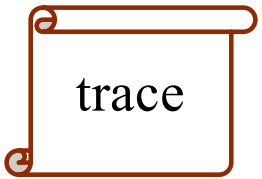
Characterization for the HA-SMR Unique Feature

- Open zone issue
  - A zone must be opened, before we write data to it. Recommended maximum #: 128
- Non-sequential written zone issue
- Media cache cleaning efficiency

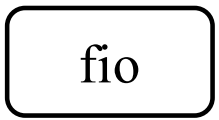
Characterization Goals

- Open zone issue
  - A zone must be opened, before we write data to it. Recommended maximum #: 128
- Under the hood:
  - Opening a zone will reserve an “open zone resource” that ensures persistence of the zone metadata through unexpected power loss.
  - When all the “open zone resources” are used up, opening a new zone will result in closing an old zone, incurring expensive disk synchronization operation.

Testing Setup

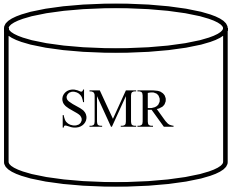


micro-benchmark trace



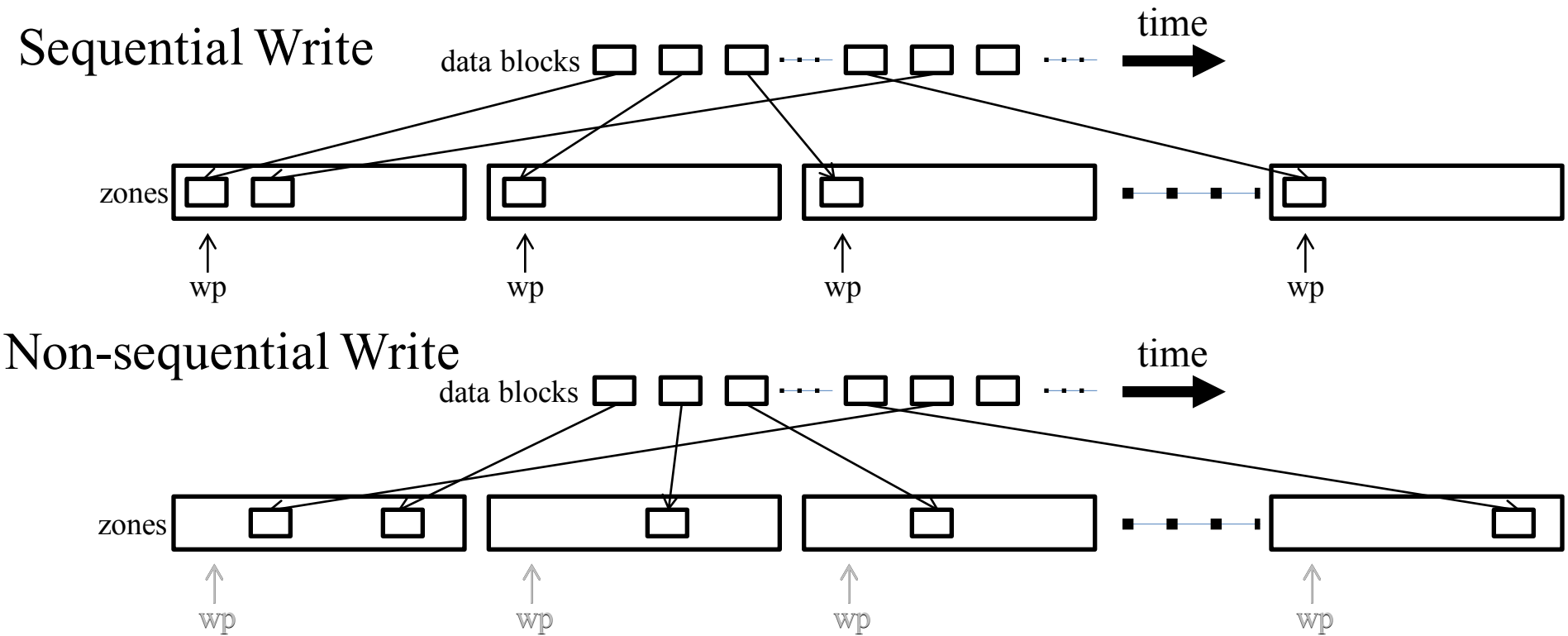
replay tool

user space ZBC library developed by HGST (branch r04)



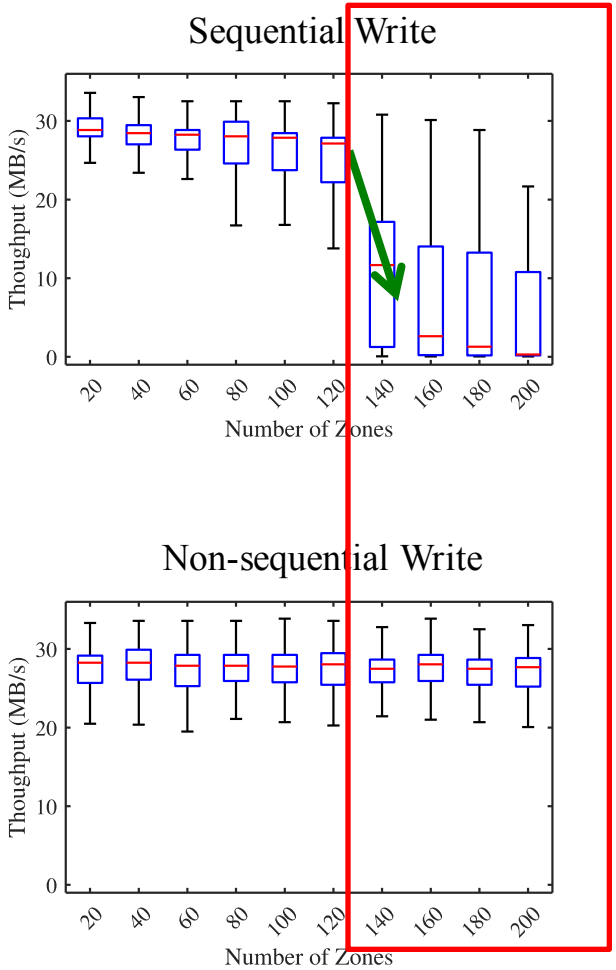
Seagate HA-SMR sample drive
model: ST8000AS0022, prototype firmware revision ZN03

Open Zone Issue Investigation



Open Zone Issues Results

- Recommended optimal number: 128.
- 1000 write requests, 4KB iosize.
- Clear performance drop for sequential write (we have proposed a solution using H-Buffer).
- No significant drop in non-sequential write.
- Sequential write does not always outperform non-sequential write.



Lesson Learned

- Designers should always respect the open zone recommendation for sequential write.
- HA-SMR does well in light-weight bursty non-sequential workload.

For other characterization results and lessons learned:
Please feel free to find me in the poster session or to refer to the paper.

Host-controlled Indirect Buffer (H-Buffer) Motivation

- Design a software layer above HA-SMR which reorganizes the workload into HA-SMR friendly ones to enhance the I/O performance.
- Exploit HA-SMR model for the design.

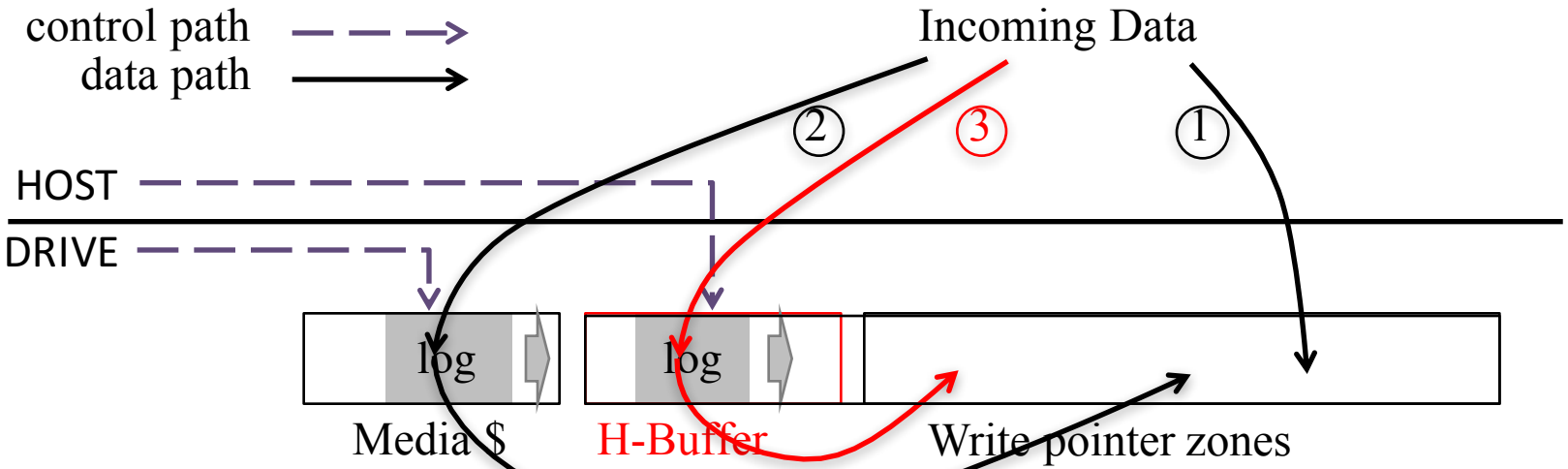
Data Handling in HA-SMR Drives

~~Host Managed~~
~~Drive Managed~~

① Sequential Write: direct path

② Non-seq Write: indirect path by **drive-controlled** media cache

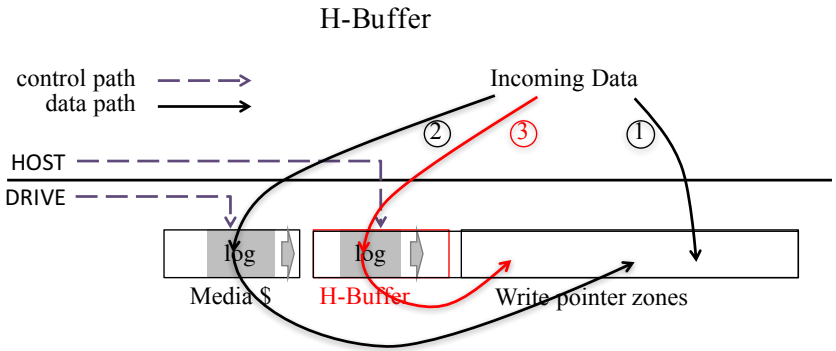
③ Seq & non-seq Write: indirect path by **host-controlled** indirection buffer

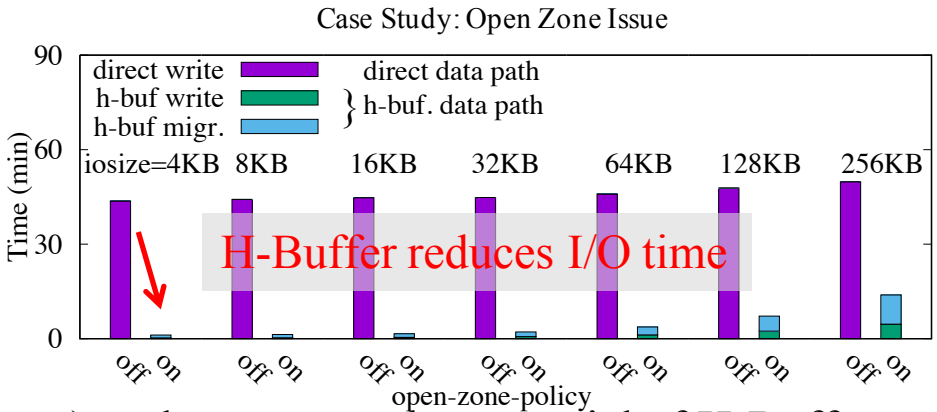


Only HA-SMR supports such a three-data-path system

Three-data-path System

- H-Buffer embodies a HA-SMR three-data-path system (the *mechanism*).
- It supports a broad spectrum of workload switching algorithm (the *policy*).
- We believe such separation of mechanism and policy can eventually lead to various solutions to the performance degradation problem.





Initial effort: a simple case study (open zone issue) to demonstrate the potential of H-Buffer.

A complete design of H-Buffer is left as future work.

Related Work

	Skylight*	Our Work
Objective	uncover the internal structure of **DM-SMR**	We focus on the unique features in **HA-SMR** (e.g. the open-zone issue), and how to meet the system design challenge by exploring an **H-Buffer** concept.
Method	software + **hardware** (high speed camera)	We leverage richer **libzbc API** to manipulate the drive and collect more information (non-sequential zone number, wp position, etc.) to aid the performance interpretation.

*A. Aghayev and P. Desnoyers, FAST'15

Summary and Future Work

- We carry out performance evaluation on HA-SMR sample drives to investigate the unique features of the HA-SMR drives.
- We summarized the system implications of the HA-SMR features and propose a novel H-Buffer that can potentially improve the performance of HA-SMR drives.
- In our future work, we plan to investigate different data path switching policies and have a complete design for H-Buffer so that HA-SMR drives can be used to construct large-scale storage systems to support various applications.

Thank You!

Fenggang Wu
fenggang@cs.umn.edu

UNIVERSITY OF MINNESOTA

Center for Research
in Intelligent Storage

National Taiwan University