SMR in Linux Systems

Seagate's Contribution to Legacy File Systems

Adrian Palmer, Drive Development Engineering

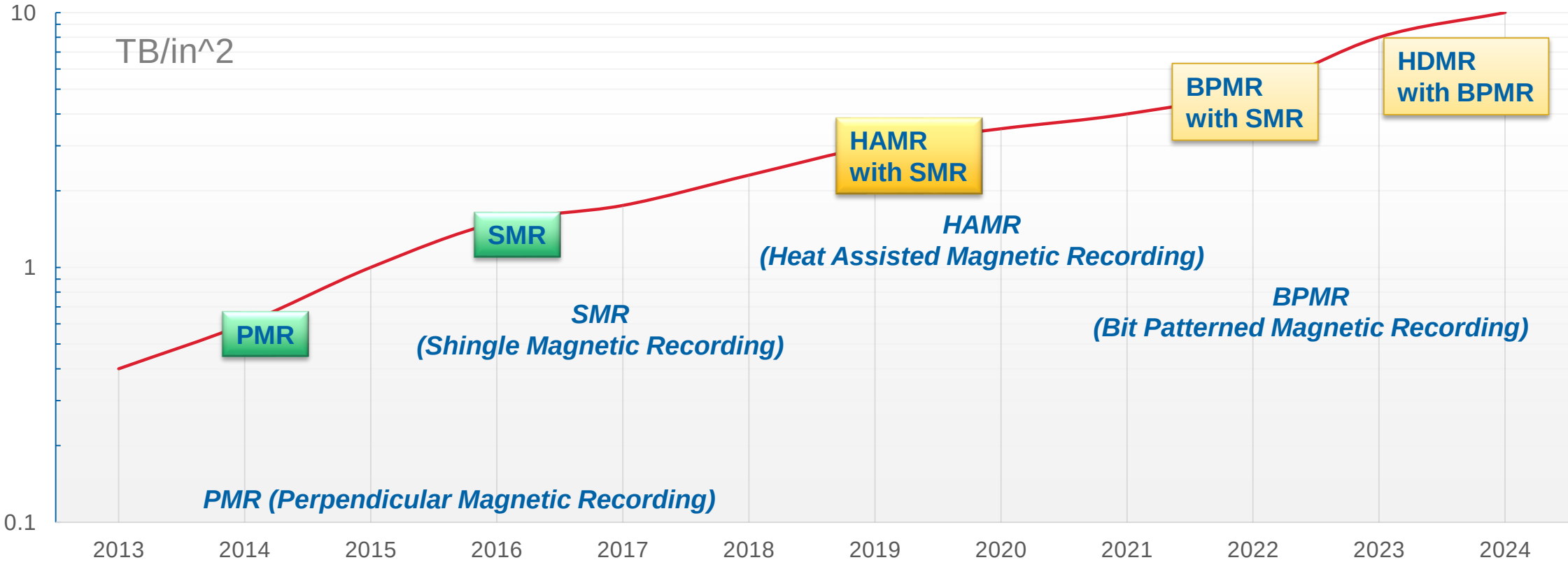
SEAGATE combines

DIFFERENT TECHNOLOGIES in new ways

to SOLVE customer data storage CHALLENGES

Shingled Magnetic Recording (SMR)

Areal density growth curve – lowest cost/GB



Data adapted from ASTC Technology Roadmap (http://www.idema.org/?page_id=5868)

Shingled Magnetic Recording (SMR)

Forward-write only: Radial AND Rotational

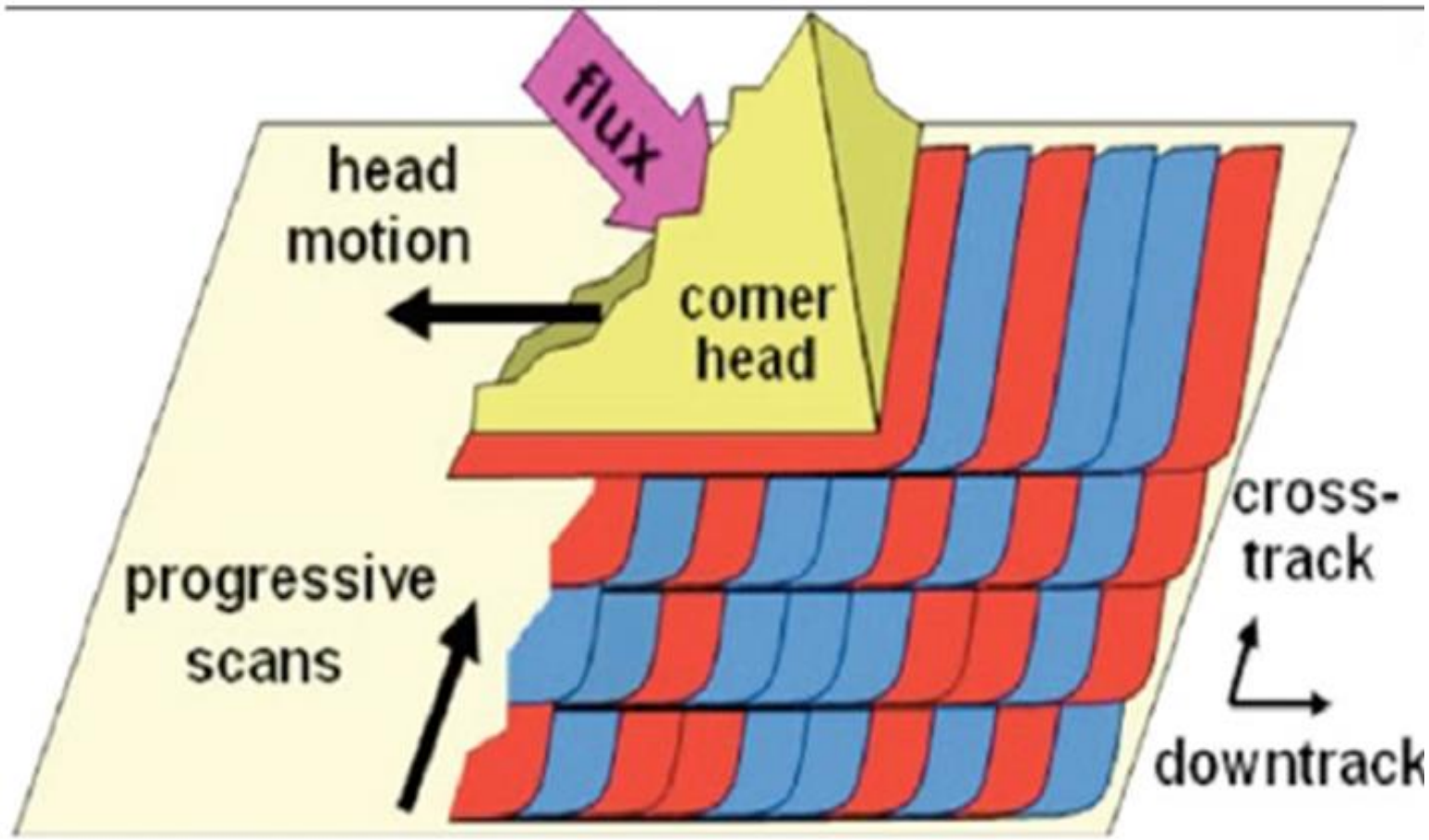
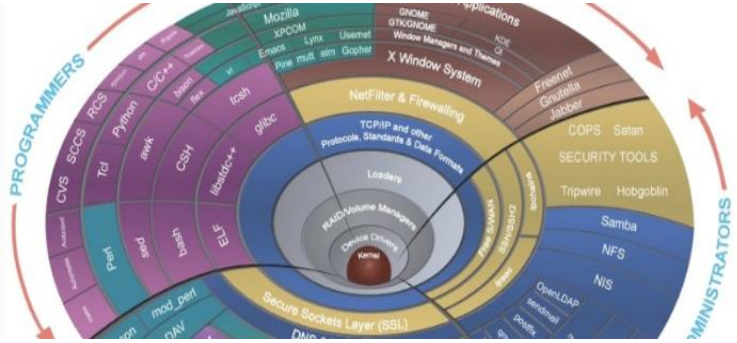


Image: Wood, R.; Williams, Mason; Kavcic, A.; Miles, Jim, "The Feasibility of Magnetic Recording at 10 Terabits Per Square Inch on Conventional Media," *Magnetics, IEEE Transactions on* , vol.45, no.2, pp.917,923, Feb. 2009

SMR Drive Types

- Drive Managed (DM)
  - Mimics Traditional drives
  - Backwards compatible
    - Direct Replacement for conventional drives in conventional apps
- Host Managed (HM)
  - Not backwards compatible
  - Host required to manage data ordering for performance mitigation
  - Extensions required in ATA and SCSI command sets
- Host Aware (HA)
  - Combination of DM and HM.
  - Backwards compatible / Able to use extensions in ATA and SCSI



The Kernel Perspective

Drive Managed

No change: regular SD drive

Host Managed

Requires new device & FS

Host Aware

Regular SD drive

FS benefits from knowledge of media layout

“

Drive Managed – Compatibility for Today

Host Aware – Performance for Tomorrow

”

Host Aware:

Capacity gains like Drive Managed

Performance like Conventional

SMR – Can we avoid it?

- Benefits
  - Provides continued growth in Areal Density.
  - Enables lower cost/GB disc drives
  - Base of new technologies – HAMR
- Support Readiness
  - ZBC/ZAC specifications are nearing completion
  - T10/T13 committees work actively progressing
- Availability
  - Millions of DM drives shipped!
  - Seagate's 8TB Archive HDD v2 drive is SMR
    - DM in production, HA forthcoming

ZBC: Zoned Block Commands
ZAC: Zoned –device ATA Commands

ZAC/ZBC Standards

- **Inspired by SMR. Applicable to any media**
- **Separates media into bounded zones**
- **Write Pointer Zones**
  - Sequential write only for Host Managed (restrictive)
  - Sequential write preferred for Host Aware (permissive)
- **New common ATA/SCSI commands**
  - REPORT ZONES
  - RESET WRITE POINTER
  - OPEN/CLOSE/FINISH ZONE
- **Requires communication with FS beyond simple Read/Write**

SMR Friendly File System

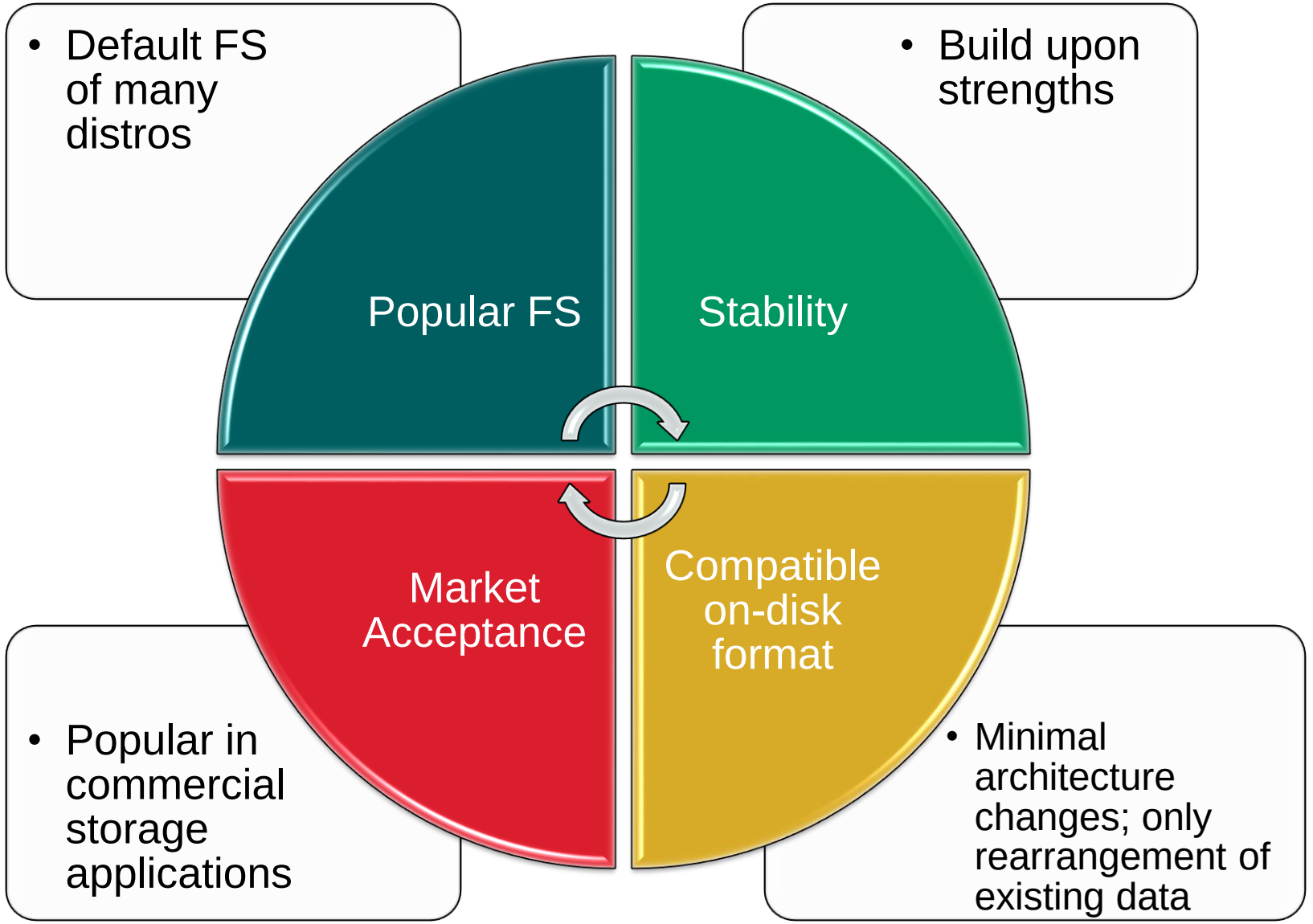
Requirements

- Forward-write only
  - CoW requirement
- Zone aware (ZAC/ZBC)
  - Boundaries
- New Commands
  - REPORT_ZONES
  - RESET_WRITE_POINTER
  - OPEN/CLOSE/FINISH ZONE
- New algorithms
  - Defragmentation

Goals

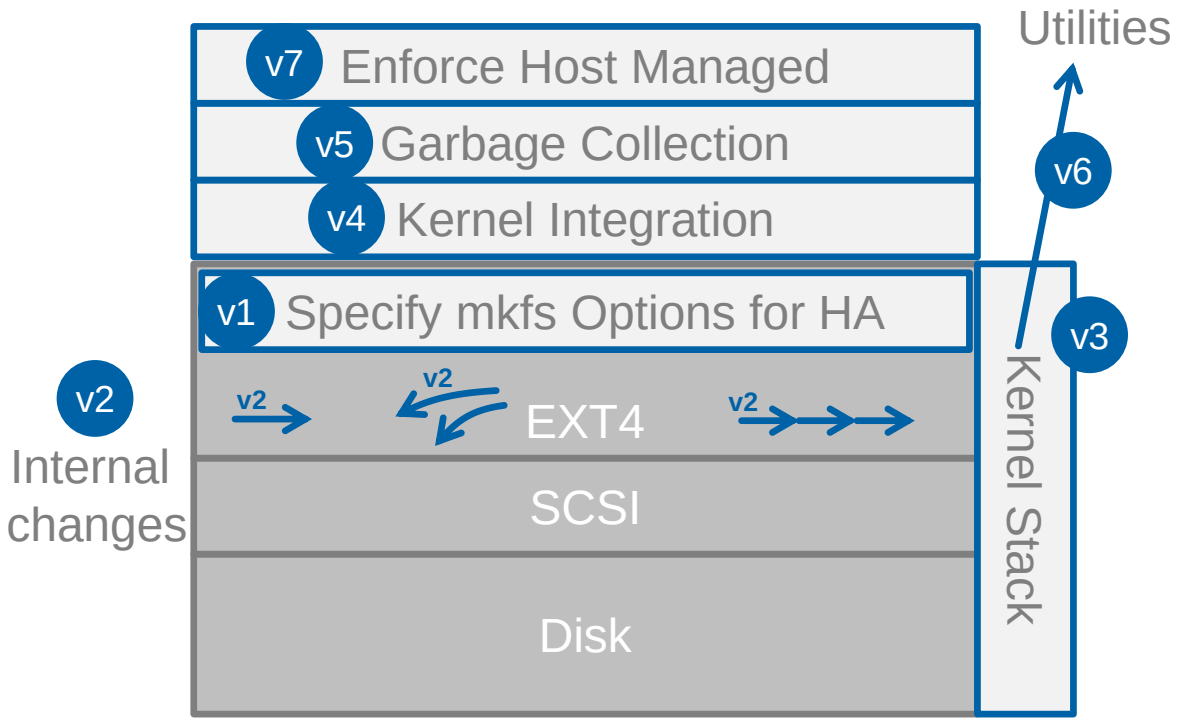
- Optimally Directed Writes – efficient streams
- Excellent Reads – streaming
- File defragmentation
- Metadata handling
- Backwards compatibility
- Provide reference design for other file systems

SMRFS -EXT4


• Default FS of many distros
• Build upon strengths
Popular FS
Stability
Market Acceptance
Compatible on-disk format
• Popular in commercial storage applications
• Minimal architecture changes; only rearrangement of existing data

Proposed Stages

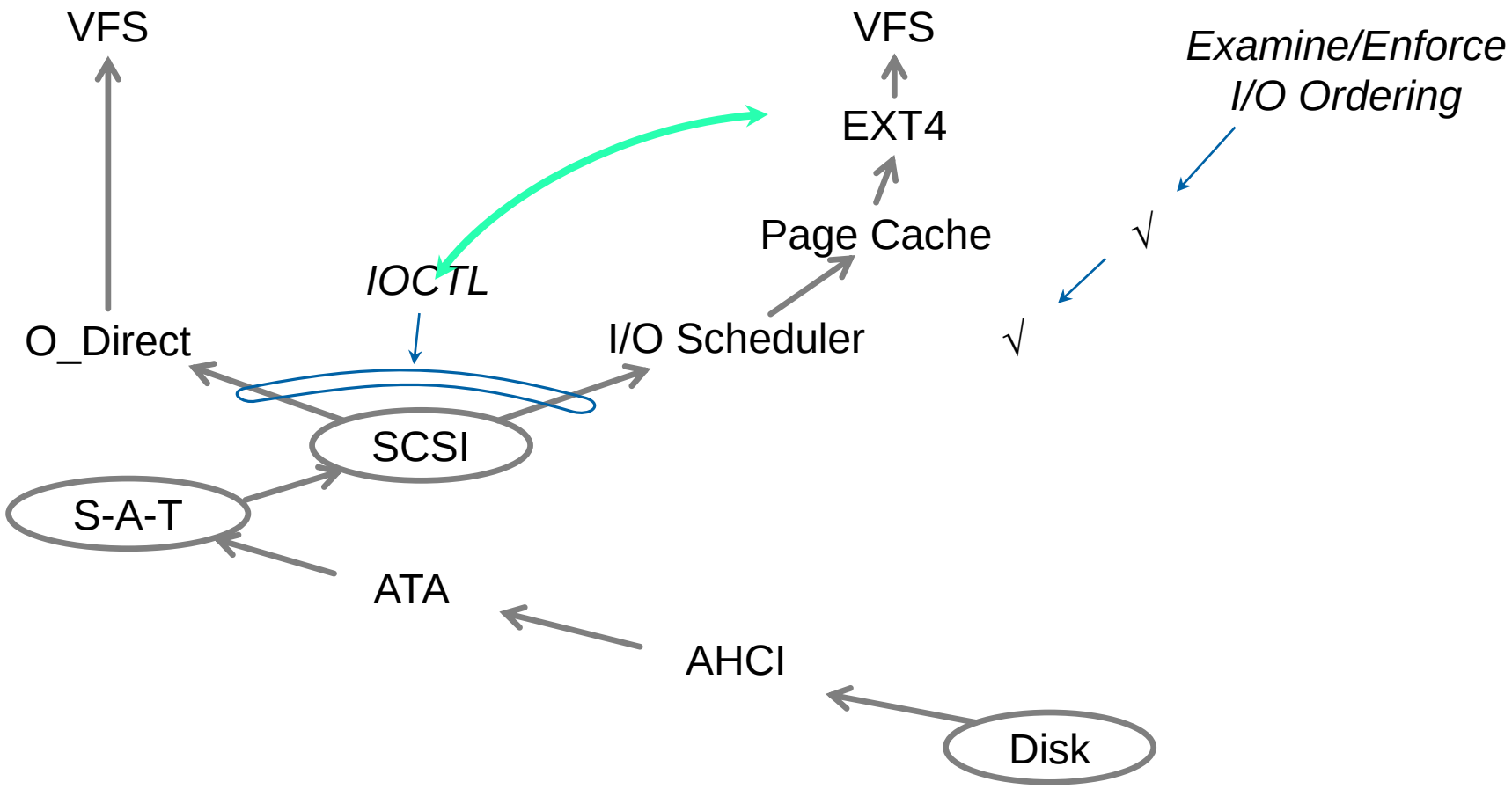
EXT4 SMRFFS Project Scope Steps



- v1 **cmd line arguments**
- v2 **internal handling changes**
- v3 **kernel stack changes**
- v4 **IOCTL integrations**
- v5 **Algorithm enhancements**
- v6 **Utility updates**
- v7 **Host Managed Compliance**

Proposed Stack Changes

V3 - Kernel Stack Changes, V4 Kernel Integration



State of project

We've done …

- Laid out design, made prototypes

Discussion at LSF…

- Consensus of key developers

We've got to do …

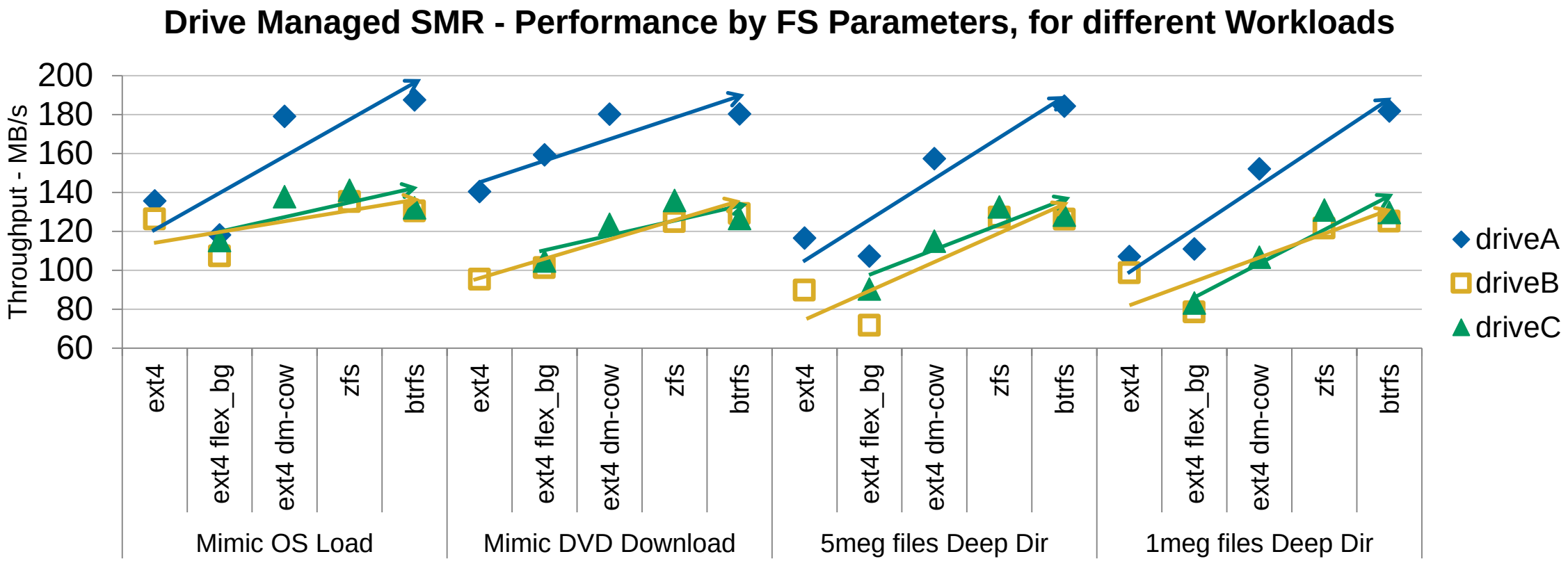
- A lot of work – in a short time

And we need community help!

- Ask how to contribute and how to get sample drives at our booth!
- Contact us after Vault at adrian.palmer@seagate.com or timothy.r.feldman@seagate.com

File System Parameters Influence Performance

For Drive Managed SMR



Re-arranging file system parameters for CoW to enforce forward-write-only improves performance of a DM-SMR-enabled system

Q&A

Thank You!

Attendees and Partners