E-Globaledge Corporation
イーグローバレッジ株式会社

# Software Defined SSD

## = HDDの特長をSSDに活かす =

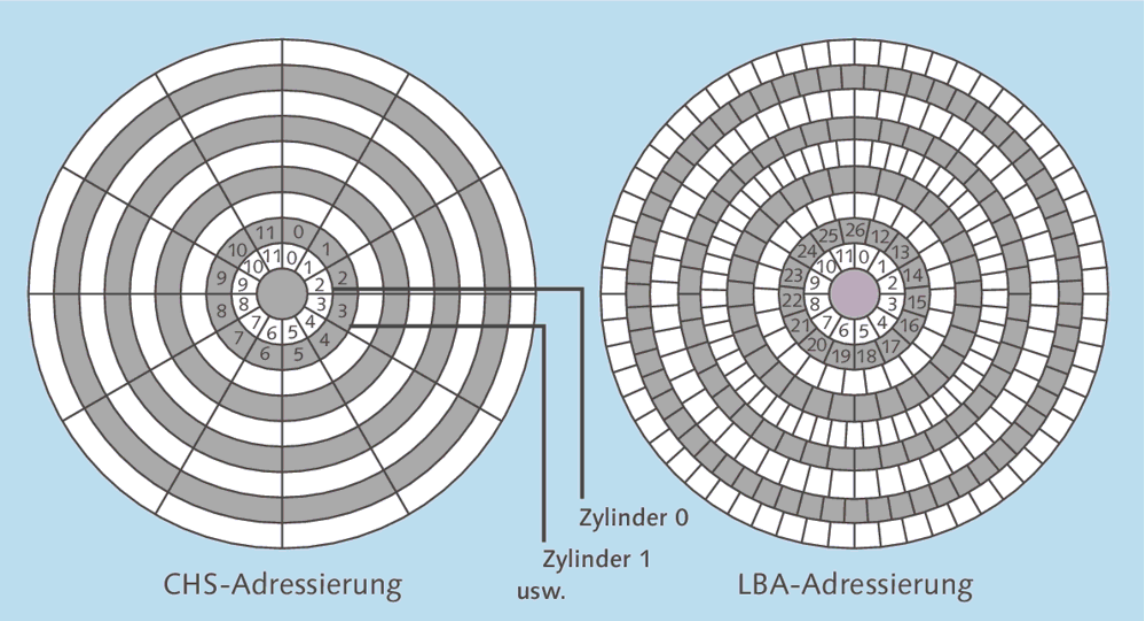
May 6, 2019
Minoru Morita
FAE, CS3 Department
E-Globaledge Corporation

# Software defined SSDは何故、必要??

# メディアのアドレス体系の歴史

E-Globaledge Corporation

| 昭和 | 平成 (現在) | 令和(近い将来) |
|---|---|---|
| CHS アドレス体系 | LBA アドレス体系 | 物理NANDアドレス体系 |
| Hostは、HDDの物理アドレスを指定してアクセス (シリンダー、ヘッド、セクタ) | Hostは、HDDの論理アドレスを指定してアクセス これは、エリアの記録密度を上げるために必要な技術 | Hostは、NANDの物理アドレスを指定してアクセス |
| 角密度一定のHDD | 線密度一定のHDD NAND Flash SSD | 超高速NAND Flash SSD |




**LBAアドレス体系はNAND Flashに根本的に合わない!!**

**主な原因として...**
**上書き対応には、複雑な制御が必須**
**セクタサイズとページサイズが不一致**

**結果、SSDの動作は非常に遅くなる**

**ゆえに新方式が必要!!**

# FTL (Flash Translation Layer)

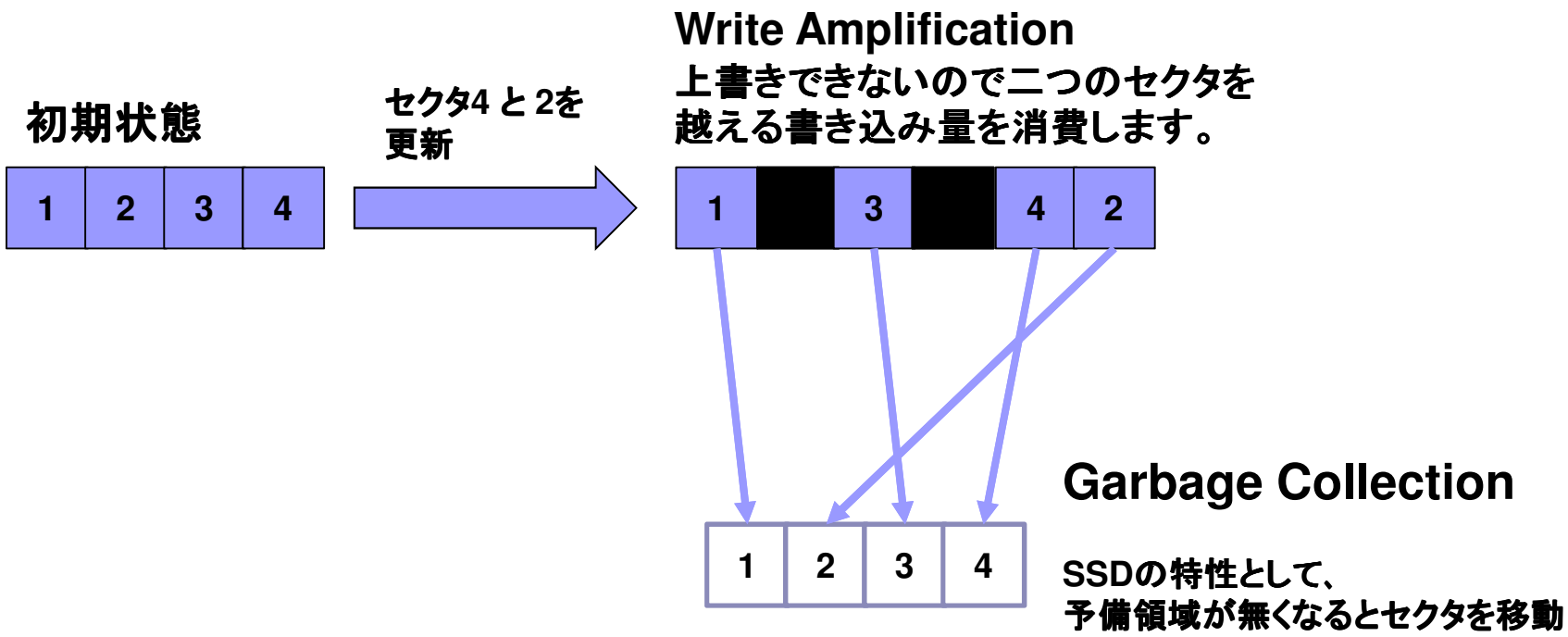


- LBA (論理アドレス)を物理NANDアドレスへ変換すること
  - LBAセクタサイズとNANDページサイズは合致せず
  - NANDフラッシュには一度にページを書き込む必要あり

# Over write vs. Copy on Write



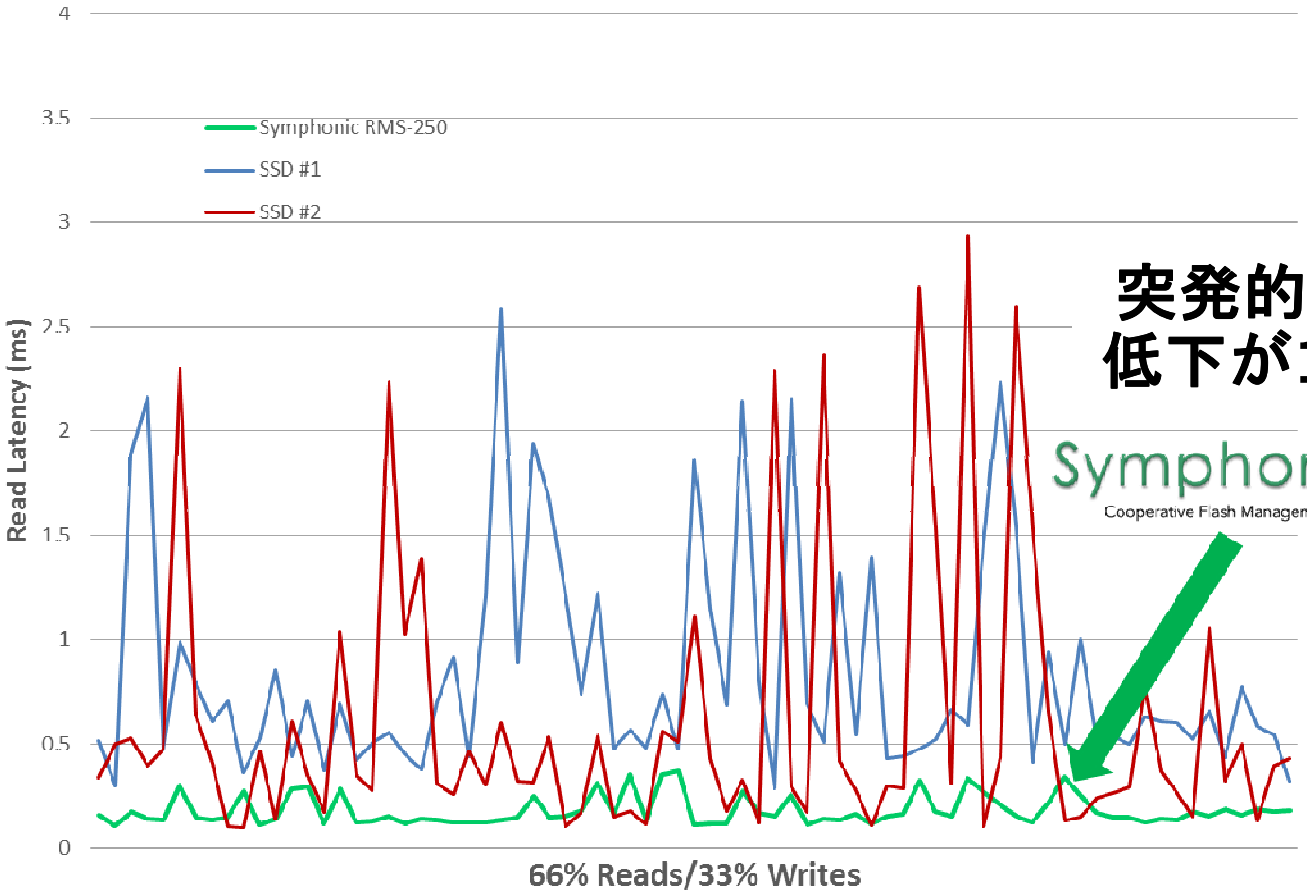
- Write Amplification (書き込み増幅度)
- Garbage Collection (ガベージコレクション)



# FTLによる反応速度低下



**ストレージアプリケーションにおいて、**
**QoS (レイテンシーとデータ処理論)は、最も重要な指標**

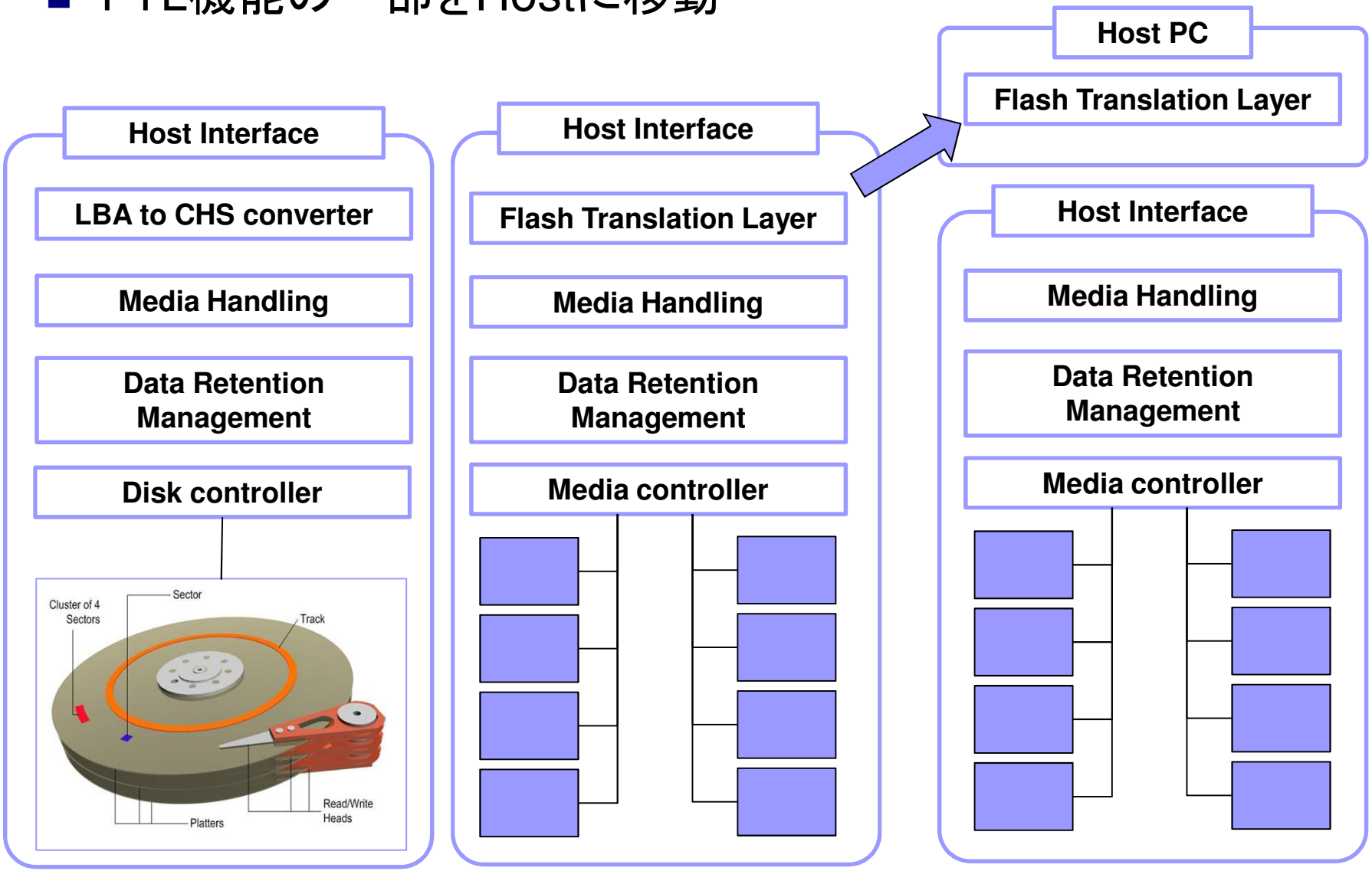



**突発的な反応速度の**
**低下が10倍以上改善**

Symphonic™
Cooperative Flash Management

# Software defined SSDの仕組み E-Globaledge Corporation

- FTL機能の一部をHostに移動



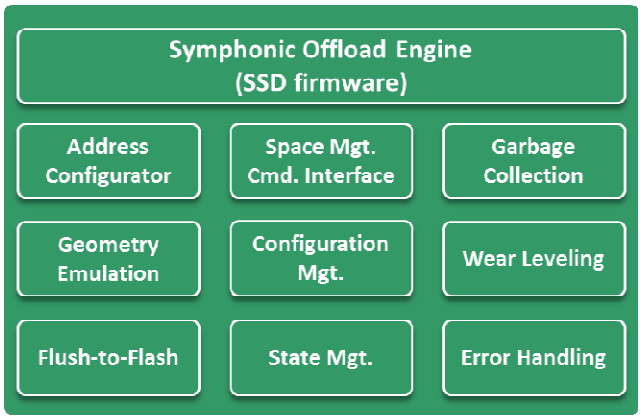
# 二つの方式が存在

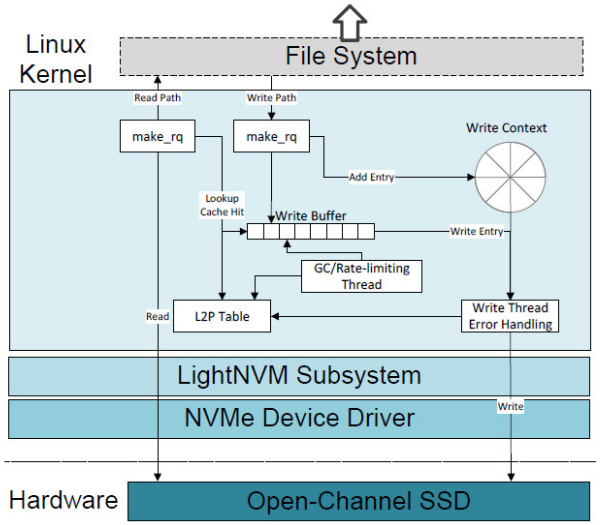
- Symphonic CFM (Radian独自)
  - Symphonic Legacy
  - Symphonic Zone
- Light NVM (Open project)

# Symphonic CFM vs. Light NVM

E-Globaledge Corporation

| Symphonic CFM<br>(CFM: Cooperative Flash Management) | Light NVM & Open-Channel |
|---|---|
| スケーラブル<br>ガベージコレクションをSSD内部で実行する為 | 不十分なスケール<br>Hostが全て管理する為、CPUやメモリを消費 |
| どんなNANDでも対応可能:<br>SSDのコントローラでNANDの特性を吸収し、<br>Hostのソフトウェアの互換性を担保<br>(Pageサイズ、Blockサイズ、セル当りのビット数) | NAND Chipの世代ごとにHostのソフトウェア開発が必要 |
| Device オフロード:<br>PCI Express バスのトラフックを増加させず、<br>SSD 内部で NAND管理を実行 | NAND管理の為、PCI Express バスのトラフィックが増加 |



# Symphonic vs. Light NVM

E-Globaledge Corporation

| | Symphonic Legacy | Symphonic Zone | Light NVM | Traditional |
|---|---|---|---|---|
| L2P Management (論理物理変換) | Cooperate with host and SSD device | SSD device Zone base FTL | Host Table is on DIMM | SSD device FTL |
| ガベージコレクション | SSD device ユーザが設定 | SSD device ユーザが設定 | Host | SSD device |
| ウェアレベリング | SSD device ユーザが設定 | SSD device ユーザが設定 | Host | SSD device |
| データリテンション | SSD device ユーザが設定 | SSD device ユーザが設定 | Host | SSD device |
| DWPD (Drive Write Per Day) Ware out management | NANDのProgram Eraseサイクルのみ依存 DWPDはユーザが管理(制御) | | | ベンダごとに定義 |

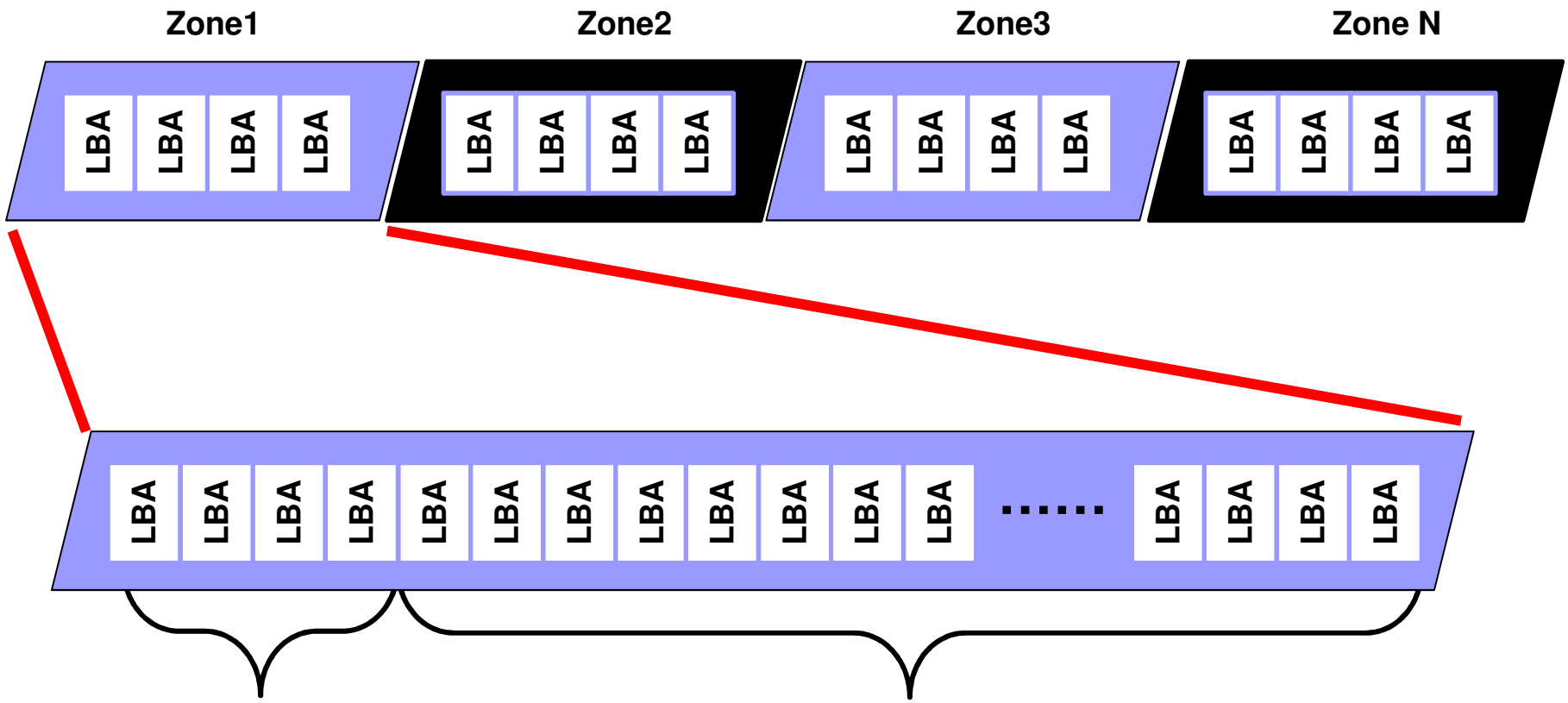
# ファームウェアのみで実現:

- 設定可能
  - ライトストライプ、ゾーン、スケジュール
- 協調型Garbage collection
- 分離型 Wear Leveling

# 設定内容



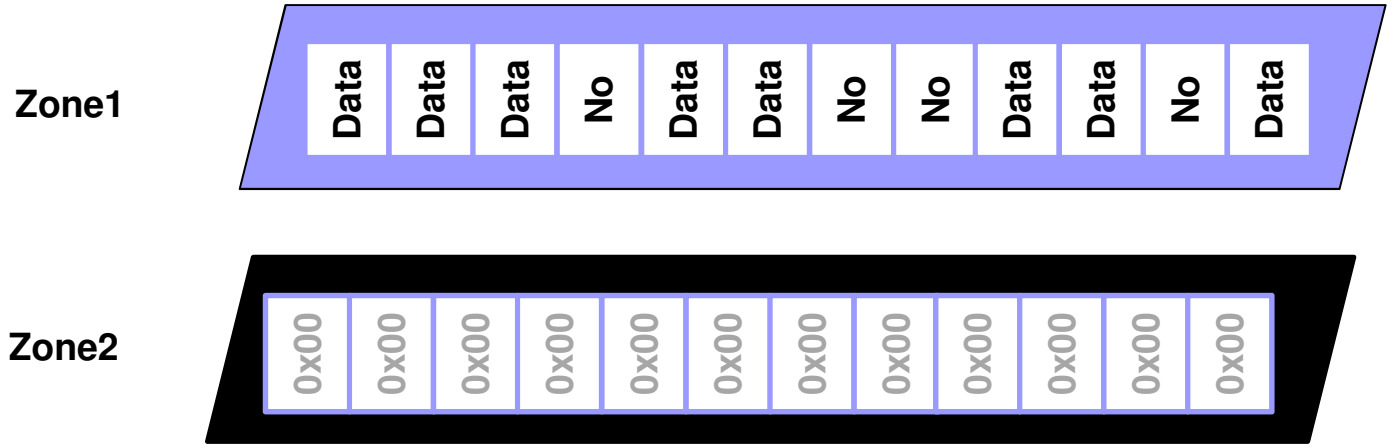
- Zone (LBAの数)
- Write stripe (論理物理テーブルの管理単位)



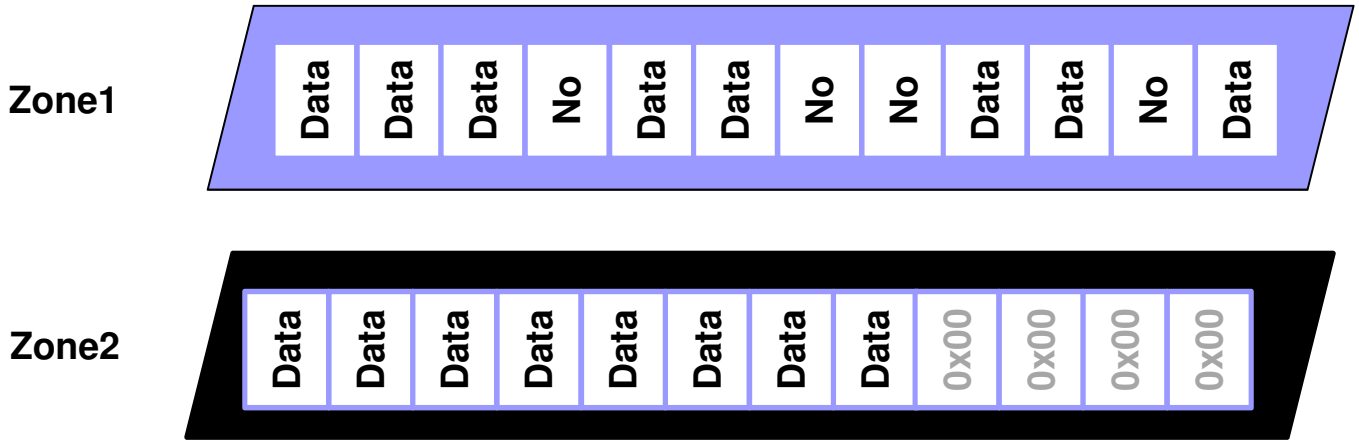
# 協調型 Garbage collection 1/2

E-Globaledge Corporation

Step1: Garbage Collection実行前



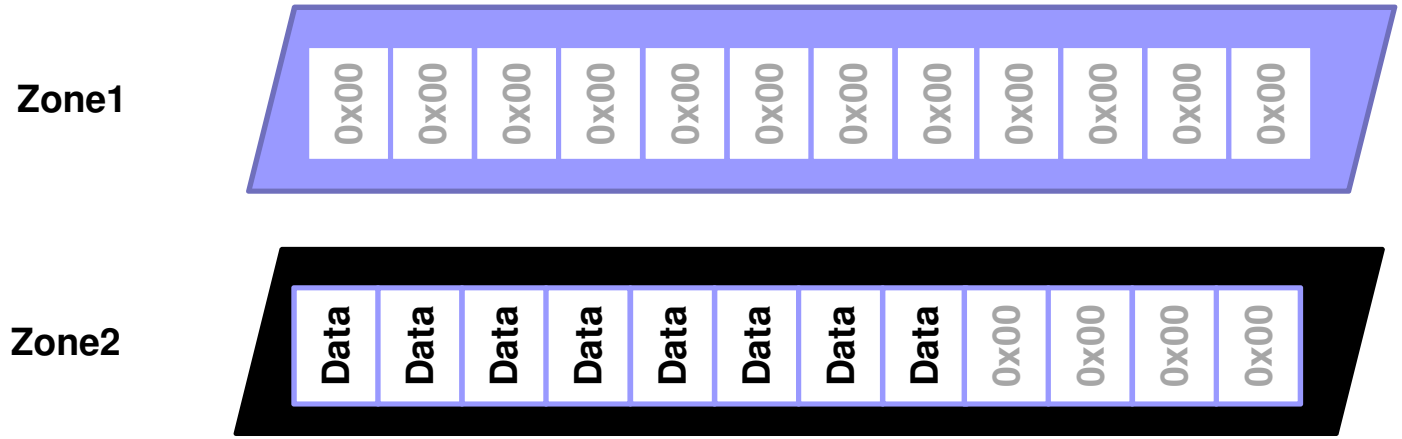

Step2: HostがZone1からZone2へ有効データのみをコピー



# 協調型 Garbage collection 2/2

E-Globaledge Corporation

Step3: HostがZone1をゼロ消去(リセット)

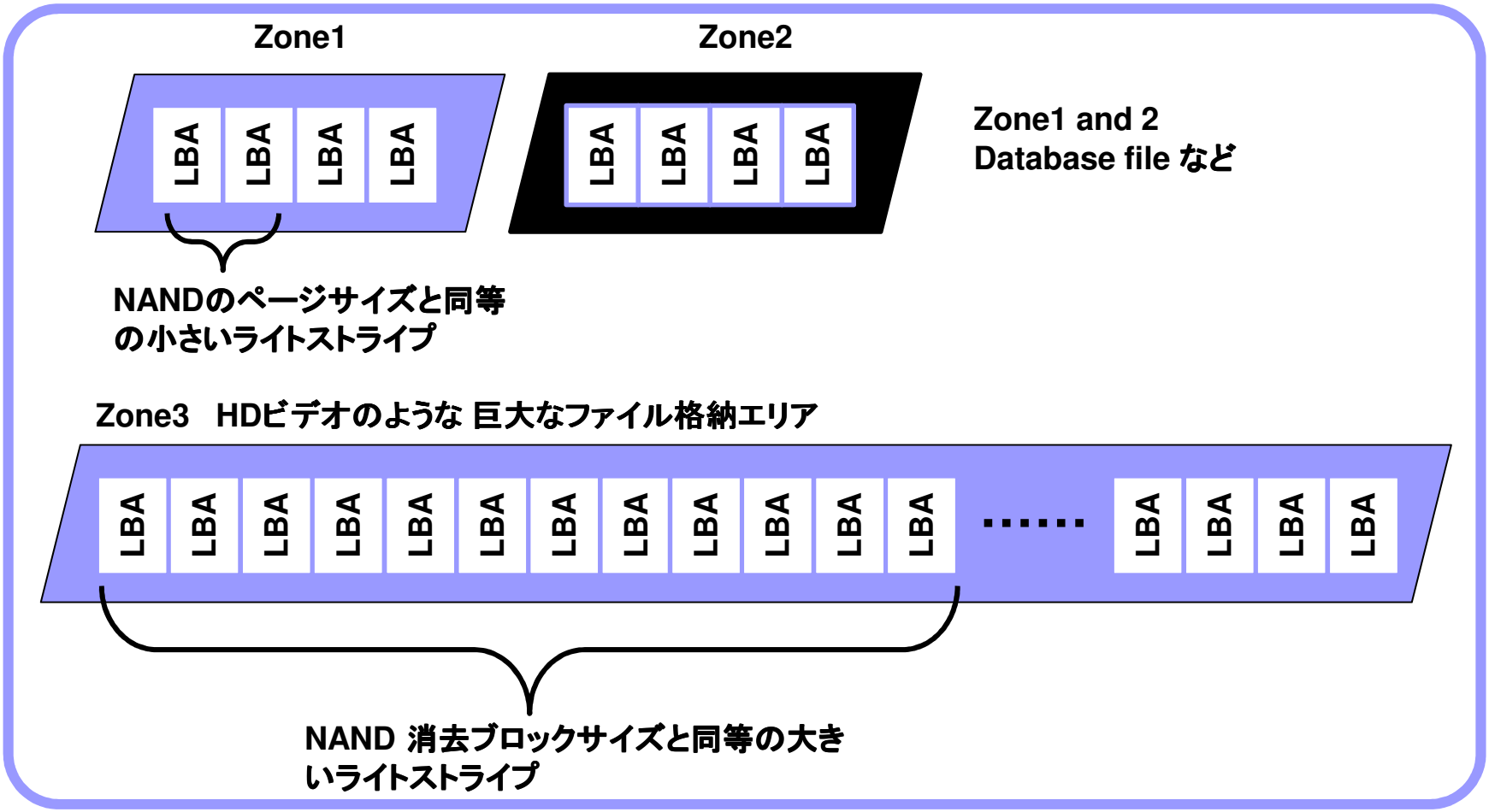



- パフォーマンス低下をコントロール出来ます
  - ユーザが把握出来ないタイミングでのガベージコレクションを実行出来なくなります
  - ガベージコレクション実行時にはウェアレベリングも実行
  - I/Oアクセスが干渉する場合はHostへ通知

# 想定使用例

E-Globaledge Corporation

- 一つのNVMe SSDをハイブリッド構成に
  - Zone1 & 2: 高IOPSが必要なアプリ用に
  - Zone 3: HDビデオ収録のような広帯域幅で巨大ファイル格納用に



# Radian社 SSD models

E-Globaledge Corporation

## RMS-325 Hybrid


- Up to 12TB eTLC
- Up to 12GB *User* NV-RAM
- PCIe x8 Gen3 NVMe interface
- DiaLog™ OEM Diagnostic Monitoring Capabilities
- Mechanism for upgrading firmware in the field

## RMS-350 Hybrid

- Up to 9TB eTLC
- Up to 12GB *User* NV-RAM
- Dual Port (2x2) or Single Port (x4)
  Gen3 NVMe interface
- 2.5” U.2 Disk Drive Format
- DiaLog™ OEM Diagnostic Monitoring Capabilities
- Mechanism for upgrading firmware in the field

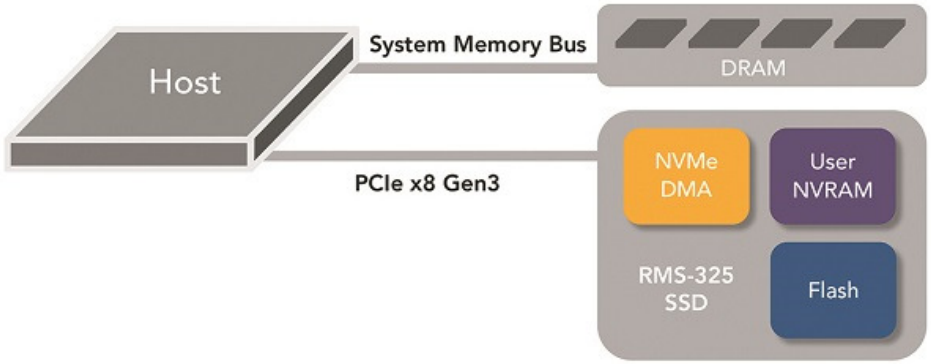
# Hybrid SSDの特別な機能

E-Globaledge Corporation

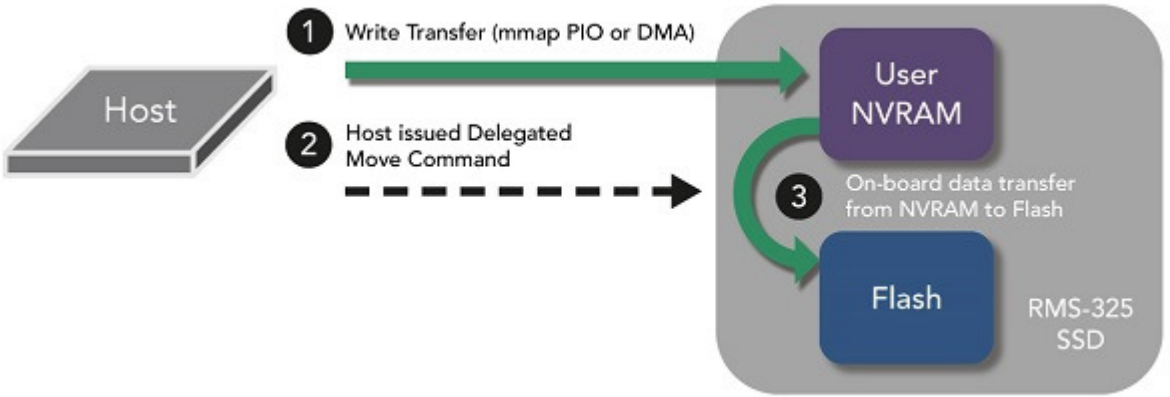
- **NV-RAM はユーザ設定可能**

**(mmap or DMA)**

- **NV-DIMM + SSDより最適化**




- **NVRAMからFlashへのドライブ内移動コマンド**
  - **Host上の CPU/Memory のリソース節約**

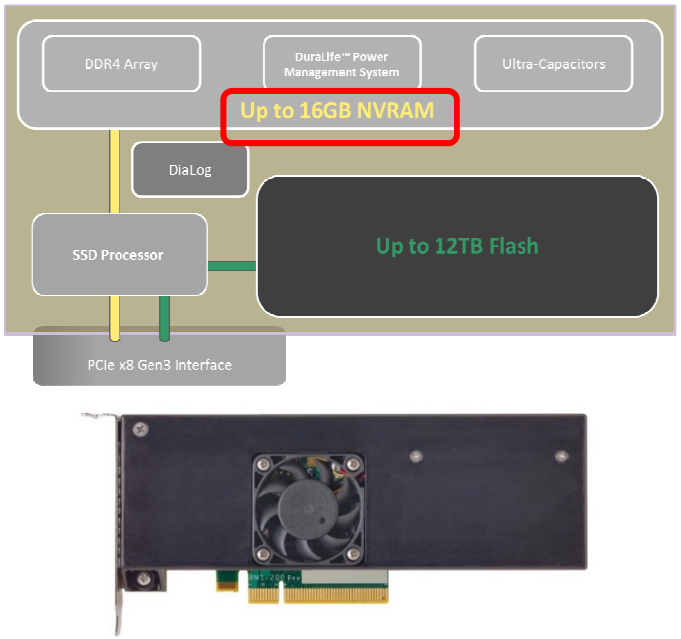


# Hybrid SSDの特長

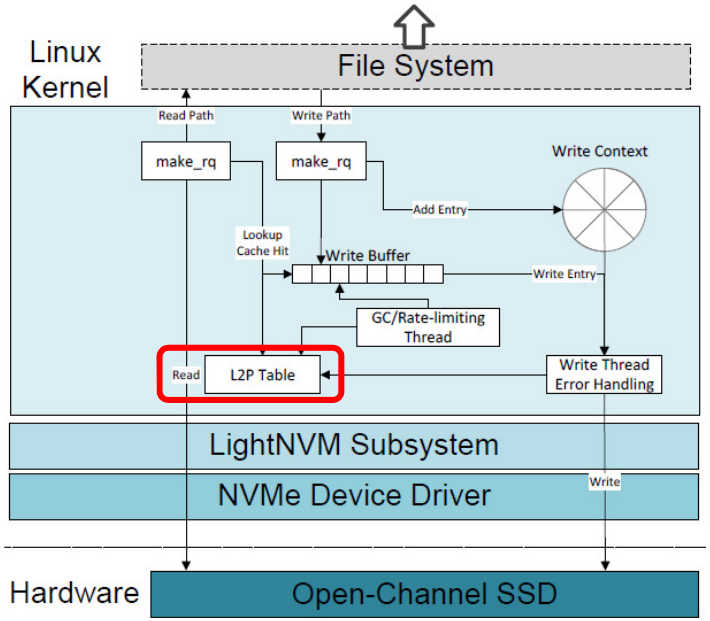
E-Globaledge Corporation

- SSD内にフラッシュ管理用NVRAMも保有
  - User NVRAM is in the SSD
  - SSD Metadata is in the NVRAM
  - Host FTL L2P (optional) in NVRAM

RMS-325 NVRAM/Flash Hybrid SSD




Other system
L2P table is in main memory (DIMM)
Volatile!!!!!!



# RAM based NVMe SSD

E-Globaledge Corporation

- RMS-375 NV-RAM

- Up to 16GB *User* NV-RAM
- Dual Port (2x2) or Single Port (x4) Gen3 NVMe interface
- 2.5” U.2 Disk Drive Format
- DiaLog™ OEM Diagnostic Monitoring Capabilities
- Mechanism for upgrading firmware in the field

- Applications: Write Caching, Journaling, Write Ahead Logging
- Lowest Latency and unlimited writes (DWPD) – unlike X-Point
- Dual Ports (2x2) ideal for ‘Dual Head’ Storage Controllers in active/active configuration
- No complex software to mirror data between nodes
- Supports Hot Swap and Live Insertion