

Flash Reliability in Production : The Expected and Unexpected

Amira Akloul, Hamza Belaoura, Maxime Kermarquer

Université de Versailles Saint-Quentin

7 janvier 2016

Introduction

- Mémoire flash est de plus en plus présente dans les centres de stockages.
- De nombreuses études sur sa fiabilité ont été réalisées en laboratoire, mais très peu sur le terrain.
- Données de monitoring fourni par Google collecté sur 6 ans qui contiennent :
 - Statistiques sur les charges de travail
 - Nombre de lectures / écritures / effacements
 - Nombre de pannes / remplacements / réparations

Introduction

- Les tests ont été réalisés sur différents disques SSD : MLC (Multiple level cell), SLC (Single level cell).

Model name	MLC-A	MLC-B	MLC-C	MLC-D	SLC-A	SLC-B	SLC-C	SLC-D	eMLC-A	eMLC-B
Generation	I	I	I	I	I	I	I	I	2	2
Vendor	I	II	I	I	I	I	III	I	I	IV
Flash type	MLC	MLC	MLC	MLC	SLC	SLC	SLC	SLC	eMLC	eMLC
Lithography (nm)	50	43	50	50	34	50	50	34	25	32
Capacity	480GB	480GB	480GB	480GB	480GB	480GB	480GB	960GB	2TB	2TB
PE cycle limit	3,000	3,000	3,000	3,000	100,000	100,000	100,000	100,000	10,000	10,000
Avg. PE cycles	730	949	529	544	860	504	457	185	607	377

Introduction

Dans cet article il y est présenté :

- Les différents types d'erreurs des opérations mémoires.
- Le métrique RBER qui évalue la fiabilité de la mémoire flash, ainsi que les différents facteurs qui l'affecte.
- Comparaisons de la fiabilité :
 - Entre les différentes technologies de mémoire flash.
 - Entre la mémoire flash (SSD) et les disques durs classiques (HDD).
- Taux de réparations / remplacements des disques.

Type d'erreurs

■ Erreurs transparentes

- Peuvent être masquées à l'utilisateur

■ Erreurs non-transparentes ou visibles

- Mènent à un échec d'une opération mémoire

Type d'erreurs

Erreurs transparentes

- **Erreurs corrigibles** : durant une opération de lecture l'erreur est détectée puis corrigée.
- **Erreur de lecture** : l'opération de lecture échoue, puis après une ou plusieurs tentatives, elle réussie.
- **Erreur d'écriture** : l'opération d'écriture échoue, puis après une ou plusieurs tentatives, elle réussie.
- **Erreur d'effacement** : l'opération d'effacement a échoué.

Type d'erreurs

Erreurs non-transparentes

- **Erreurs incorrigibles** : Une opération de lecture compte plus de bits corrompus que l'on puisse corriger.
- **Erreur de lecture finale** : l'opération de lecture échoue, même après plusieurs tentatives.
- **Erreur d'écriture finale** : l'opération d'écriture échoue, même après plusieurs tentatives.
- **Erreur de méta-données** : l'opération d'accès aux méta-données échoue.
- **Erreur de délai** : le temps de l'opération mémoire est supérieure à 3 secondes.

Fréquences des erreurs

- L'erreur de lecture finale est la plus courante, deux fois plus fréquente que tous les autres types d'erreurs non-transparentes.
- Les erreurs de méta-données et d'écriture finales ont des fréquences proches, très faible par rapport à la fréquence d'erreur de lecture finale.
- L'erreur corrigible est la plus courante des erreurs transparentes, suivie par les erreurs d'écriture et les erreurs d'effacement.

⇒ Les erreurs transparentes sont plus rares que les erreurs non-transparentes.

Le métrique RBER

- Raw Bit Error Rates (RBER) est la mesure standard pour évaluer la fiabilité du disque.
- Elle est définie comme :

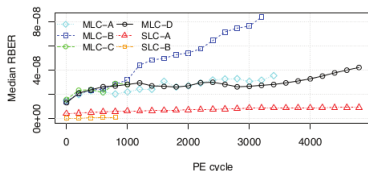
$$\frac{\text{Nombre de bits corrompus}}{\text{Nombre total de bits lus}}$$

- Plusieurs facteurs impactent le RBER :
 - L'usure (wear-out)
 - L'âge
 - La lithographie
 - La charge de travail (workload)

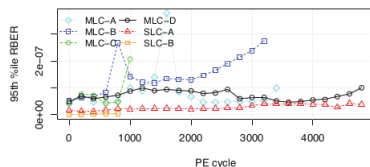
Facteurs impactant le RBER

■ Wear-out (l'usure)

- 1 RBER en fonction du wear-out : Une croissance linéaire.
- 2 Le taux de croissance est différent selon les modèles.
- 3 Les tests en laboratoire diffèrent des tests sur le terrain. Les disques testés en laboratoire de façon intensive, et ont une usure (PE cycles limit) plus limitée.



(a)

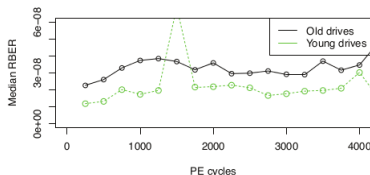


(b)

Figure 2: The figures show the median and the 95th percentile RBER as a function of the program erase (PE) cycles.

Facteurs impactant le RBER

- **Age** : Le nombre de mois qu'un dispositif est sur le terrain.



⇒ Plus le disque est ancien plus il fait d'erreurs.

- **Lithographie** Une petite lithographie implique un grand taux de RBER.

Facteurs impactant le RBER

- **Workload (charge de travail)** Type et nombre d'opérations réalisées par le disque. Les corruptions de bits sont causées par :
 - Erreur de rétention (perte de la charge).
 - Perturbation par une lecture d'une cellule voisine.
 - Perturbation par une écriture dans une cellule voisine.
 - Effacement incomplet de la cellule.
- ⇒ Les lectures perturbatrices affectent beaucoup plus le taux de RBER pour certains disques.
- ⇒ Pas de preuve que les perturbations par écriture et les effacements incomplets ont un impact sur le RBER.

Facteurs impactant le RBER

⇒ Il est difficile de prédire les taux de RBER en se basant sur les tests faits en laboratoire. Les taux de RBER prédits par les tests de laboratoire sont différents des taux qu'on observe sur le terrain.

Problèmes matérielles

- **Bad blocks** : Un bad block est une case mémoire ou une erreur de lecture finale, d'écriture, d'effacement a été détectée.
- **Bad chips** Une puce qui a plus de 5% de de bad blocks. Les disques peuvent tolérer ces pannes de deux façons :
 - Utilisation d'une puce de rechange lorsqu'ils en ont.
 - On retire la puce affectée et le disque fonctionne avec une capacité réduite.
- Durant les études il a été constaté que 2/3 de bad chips n'ont pas respecté les spécifications du vendeur et dépassent 5% des blocks échoués (Alors qu'ils sont censés ne pas dépasser plus de 2% de blocks echoués).

⇒ Si on a 2 à 4 bad blocks sur une puce on aura 50% de chance d'avoir de nouveaux bad blocks.

Réparations et remplacements

Model name	MLC-A	MLC-B	MLC-C	MLC-D	SLC-A	SLC-B	SLC-C	SLC-D	eMLC-A	eMLC-B
Drives w/ bad chips (%)	5.6	6.5	6.6	4.2	3.8	2.3	1.2	2.5	1.4	1.6
Drives w/ repair (%)	8.8	17.1	8.5	14.6	9.95	30.8	25.7	8.35	10.9	6.2
MTBRepair (days)	13,262	6,134	12,970	5,464	11,402	2,364	2,659	8,547	8,547	14,492
Drives replaced (%)	4.16	9.82	4.14	6.21	5.02	10.31	5.08	5.55	4.37	3.78

- Les disques flash SLC-C et SLC-B sont partis en réparation plus souvent que les autres disques.
- 96% des disques flash sont partis au moins une fois en réparation lors de leur exploitation.
- MLC-B et SLC-B sont à 10% de taux de remplacement qui est plus que les autres modèles qui tournent autour de 5% de taux de remplacement.

Comparaison entre technologies de mémoires flash

- Les disques flash SLC ont des taux d'erreur plus faibles que les autres modèles mais sont remplacés plus souvent. Ceci ne les rends pas plus fiables que les MLC ou les autres modèles.
- Les disques eMLC ont des taux de RBER plus grand que les MLC, dû à leur lithographie plus petite.

Comparaison entre les SSD (disques flash) et HDD (disque dur classique)

- Les disques flash ont un taux de remplacement plus faible que les disques durs classiques.
- Les SSD ont des taux d'erreurs incorrigibles plus grands que les HDD.

Conclusion

- Le taux RBER en fonction de l'usure croît de façon linéaire et non exponentielle.
- Les disques SLC, considérés par les entreprises comme meilleurs, ne sont pas plus fiables que les disques MLC.
- Sur 4 ans, 30-80% des disques ont eu au moins un bad block, et 2-7% au moins un bad chip.

Merci de votre attention.