

# Table des matières

<b>La série <i>Gestion de l'énergie dans les systèmes embarqués</i></b> . . . . .	13
<b>Remerciements</b> . . . . .	15
<b>Préface</b> . . . . .	17
Pierre FICHEUX	
<b>Partie 1. Introduction</b> . . . . .	19
<b>Chapitre 1. Introduction générale</b> . . . . .	21
1.1. De l'explosion des données numériques . . . . .	21
1.2. Performance et consommation des systèmes de stockage . . . . .	22
1.3. Hiérarchie mémoire et technologie de stockage . . . . .	23
1.4. Introduction à l'intégration des mémoires flash . . . . .	26
1.5. Objectif du livre . . . . .	28
1.6. Publics visés . . . . .	28
1.7. Plan du livre . . . . .	29
1.8. Comment lire ce livre . . . . .	30
<b>Chapitre 2. Mémoires flash : structure et contraintes</b> . . . . .	31
2.1. Présentation générale des mémoires flash . . . . .	31
2.1.1. Différents types de mémoires flash . . . . .	31
2.1.2. Principes physiques de fonctionnement . . . . .	32

2.1.3. Architecture hiérarchique simplifiée d'une puce de mémoire flash NAND	33
2.1.3.1. Page, bloc et plan	34
2.1.3.2. Notion de dé ( <i>die</i> ) et d'unité logique ou LUN ( <i>Logical Unit</i> )	35
2.1.3.3. Structure des composants de stockage au sein d'un SSD	36
2.1.4. Opérations sur les mémoires flash	37
2.1.4.1. Opérations de base	37
2.1.4.2. Opérations avancées	38
2.1.4.3. Latences et consommation des opérations flash	39
2.2. Contraintes et limitations	40
2.2.1. Contrainte d'effacement avant écriture	40
2.2.2. Contrainte d'usure	40
2.2.3. Contrainte de fiabilité	41
2.3. Systèmes de gestion de contraintes : concepts généraux	41
2.3.1. Gestion de la contrainte d'effacement avant écriture	42
2.3.1.1. Traduction d'adresses logiques vers physiques	42
2.3.1.2. Ramasse-miettes	43
2.3.2. Répartition de l'usure	45
2.3.3. Gestion de la fiabilité	46
2.3.4. Systèmes de gestion de contraintes	47
2.4. Conclusion	47

### Chapitre 3. Evaluation de performance et de consommation énergétique des systèmes de stockage

3.1. Benchmarking pour les systèmes de stockage à base de mémoire flash	49
3.1.1. Microbenchmarks	51
3.1.1.1. Microbenchmarks niveau bloc	51
3.1.1.2. Microbenchmarks niveau système de fichiers	52
3.1.2. Macrobenchmarks	53
3.1.2.1. Autres macrobenchmarks pour systèmes de stockage	54
3.1.3. Traces d'E/S	54
3.1.3.1. Traces d'E/S fréquemment utilisées dans le domaine du stockage	54
3.1.3.2. Récolte des traces d'E/S	55
3.2. Métriques de performance et de consommation des systèmes de stockage	55
3.2.1. Métriques de performance	55
3.2.1.1. Métriques d'évaluation des performances pures	56

3.2.1.2. Métriques d'évaluation de la répartition de l'usure . . . . .	56
3.2.1.3. Métriques d'évaluation propres aux FFS . . . . .	57
3.2.2. Métriques de consommation énergétique . . . . .	57
3.3. Mesures de performance et de consommation des systèmes de stockage à base de mémoire flash . . . . .	58
3.3.1. Exploration des performances par la mesure . . . . .	58
3.3.1.1. Mesures de temps . . . . .	58
3.3.1.2. Mesures d'occurrence d'événements et autres métriques . . . . .	59
3.3.2. Exploration de la consommation énergétique des systèmes de stockage à base de mémoire flash NAND par la mesure . . . . .	60
3.4. Evaluation des performances et de la consommation par simulation . . . . .	60
3.4.1. Notions générales . . . . .	60
3.4.2. FlashSim . . . . .	62
3.4.3. NandFlashSim . . . . .	63
3.4.4. Autres simulateurs et comparaison générale . . . . .	63
3.5. Conclusion . . . . .	65

## **Partie 2. Domaine de l'embarqué et systèmes de fichiers dédiés aux mémoires flash : *Flash File Systems*** . . . . . 67

### **Chapitre 4. Systèmes de fichiers dédiés aux mémoires flash** . . . . . 69

4.1. Présentation générale des FFS . . . . .	69
4.1.1. Gestion du stockage et contraintes des mémoires flash . . . . .	70
4.1.2. Contraintes embarquées et passage à l'échelle (scalabilité) des FFS . . . . .	71
4.2. Intégration des systèmes de stockage de type FFS dans les systèmes informatiques : l'exemple de Linux . . . . .	72
4.2.1. Le système de fichiers virtuel de Linux : VFS . . . . .	73
4.2.2. Le pilote NAND MTD . . . . .	73
4.3. Présentation des FFS les plus populaires : JFFS2, YAFFS2 et UBIFS . . . . .	74
4.3.1. JFFS2 . . . . .	75
4.3.1.1. JFFS2 : concepts généraux . . . . .	75
4.3.1.2. JFFS2 : indexation et problèmes de passage à l'échelle . . . . .	76
4.3.1.3. JFFS2 : ramasse-miettes et autres fonctionnalités . . . . .	77
4.3.2. YAFFS2 . . . . .	78
4.3.2.1. YAFFS2 : fonctionnement général . . . . .	78
4.3.2.2. YAFFS2 : indexation, ramasse-miettes, répartition de l'usure . . . . .	79

4.3.3. UBI et UBIFS . . . . .	80
4.3.3.1. UBIFS : la couche UBI . . . . .	80
4.3.3.2. UBIFS : fonctionnement général et indexation . . . . .	80
4.3.3.3. UBIFS : ramasse-miettes et répartition de l'usure . . . . .	81
4.4. Etat de l'art des autres FFS . . . . .	82
4.5. Conclusion . . . . .	84

## **Chapitre 5. Méthodologie d'exploration de performance et de consommation énergétique pour systèmes de fichiers dédiés aux mémoires flash . . . . .**

85

5.1. Présentation générale de la méthodologie d'exploration . . . . .	86
5.1.1. Méthode et outils . . . . .	86
5.1.2. Plateforme matérielle . . . . .	87
5.2. Une suite d'outils pour l'exploration des performances des systèmes à base de FFS sous Linux . . . . .	88
5.2.1. Flashmon . . . . .	89
5.2.1.1. Flashmon : concepts de base . . . . .	89
5.2.1.2. Flashmon : fonctions sondées . . . . .	91
5.2.1.3. Flashmon : sorties et contrôle de la trace . . . . .	92
5.2.1.4. Flashmon : conclusion . . . . .	93
5.2.2. VFSSMon et FuncMon . . . . .	93
5.2.2.1. VFSSMon . . . . .	93
5.2.2.2. FuncMon . . . . .	95
5.2.2.3. Interopérabilité des outils de trace Flashmon, VFSSMon et FuncMon . . . . .	95
5.3. Exploration de la consommation énergétique : la plateforme Open-PEOPLE . . . . .	96
5.4. Conclusion . . . . .	99

## **Chapitre 6. Performance et consommation des systèmes de fichiers dédiés : résultats expérimentaux . . . . .**

101

6.1. Niveau matériel et pilote . . . . .	101
6.1.1. Performance et consommation des opérations flash de base . . . . .	102
6.1.2. Impact du tampon de lecture MTD . . . . .	104
6.1.3. Conclusion pour le niveau matériel et pilote . . . . .	106
6.2. Niveau FFS : focus sur JFFS2 . . . . .	106
6.2.1. Lecture au niveau FFS . . . . .	107
6.2.1.1. Notions générales . . . . .	107
6.2.1.2. Impact de la fragmentation . . . . .	112

6.2.2. Ecriture au niveau FFS . . . . .	114
6.2.2.1. Notions générales . . . . .	114
6.2.2.2. Ramasse-miettes et état initial . . . . .	118
6.2.3. Conclusion de l'exploration au niveau FFS . . . . .	122
6.3. Niveau VFS . . . . .	122
6.3.1. Page cache . . . . .	122
6.3.2. Le mécanisme de lecture anticipée : <i>read-ahead</i> . . . . .	124
6.3.3. Le mécanisme d'écriture retardée : <i>write-back</i> . . . . .	126
6.3.4. Conclusion au niveau VFS . . . . .	128
6.4. Conclusion . . . . .	128

### Partie 3. Les couches de traduction flash :

<b>Flash Translation Layer . . . . .</b>	<b>131</b>
--	------------

### Chapitre 7. Présentation des couches de traduction

<b>des mémoires flash . . . . .</b>	<b>133</b>
7.1. Introduction . . . . .	133
7.2. Mécanismes de base de traduction d'adresse . . . . .	134
7.2.1. Mécanisme de traduction par page . . . . .	134
7.2.2. Mécanisme de traduction par bloc . . . . .	135
7.2.3. Mécanisme de traduction hybride . . . . .	136
7.3. Mécanismes de traduction d'adresse complexes . . . . .	136
7.3.1. FTL à bloc de log . . . . .	137
7.3.1.1. Mitsubishi . . . . .	138
7.3.1.2. M-Systems . . . . .	138
7.3.1.3. BAST ou <i>Block Associative Sector Translation FTL</i> . . . . .	139
7.3.1.4. FAST ou <i>Fully Associative Sector Translation FTL</i> . . . . .	139
7.3.1.5. LAST ou <i>Locality Aware Sector Translation FTL</i> . . . . .	140
7.3.2. FTL à traduction par page . . . . .	141
7.3.2.1. DFTL ou <i>Demand-Based FTL</i> . . . . .	141
7.3.2.2. SFTL ou <i>Spatial Locality Aware FTL</i> . . . . .	142
7.3.2.3. CFTL ou <i>Convertible FTL</i> . . . . .	142
7.3.3. FTL avec mémoire flash partitionnée . . . . .	142
7.3.3.1. WAFTL ou <i>Workload Aware FTL</i> . . . . .	142
7.3.3.2. CACH-FTL ou <i>Cache Aware Configurable Hybrid FTL</i> . . . . .	143
7.4. La répartition de l'usure ou <i>wear levelling</i> . . . . .	143
7.4.1. Répartition de l'usure basée sur le nombre d'effacements . . . . .	144
7.4.2. Répartition de l'usure basée sur le nombre d'écritures . . . . .	144
7.5. Algorithmes de ramasse-miettes . . . . .	145
7.6. Mécanismes de cache pour mémoire flash . . . . .	146
7.7. Conclusion . . . . .	148

<b>Chapitre 8. Méthodologie d'évaluation de performance et de consommation énergétique des SSD</b> . . . . .	<b>151</b>
8.1. Introduction . . . . .	152
8.1.1. Méthode et outils . . . . .	152
8.1.2. Plateforme matérielle . . . . .	153
8.2. Pile logicielle d'E/S sous Linux . . . . .	153
8.2.1. La couche bloc générique . . . . .	154
8.2.2. L'ordonnanceur d'E/S . . . . .	155
8.2.3. Le pilote de périphérique . . . . .	157
8.3. Le cadre de travail : le Cloud . . . . .	157
8.4. Outils de trace d'E/S pour l'exploration des performances . . . . .	158
8.4.1. Niveau 1 : l'hyperviseur . . . . .	159
8.4.2. Niveau 2 : VFS de l'hôte . . . . .	161
8.4.2.1. Utilisation de l'outil <i>strace</i> . . . . .	161
8.4.2.2. Utilisation d'un module avec <i>Jprobe(s)</i> . . . . .	162
8.4.3. Niveau 3 : FS . . . . .	163
8.4.4. Niveau 4 : bloc . . . . .	163
8.4.4.1. Utilisation de <i>blktrace</i> . . . . .	164
8.4.4.2. Utilisation d'un module avec <i>Jprobe(s)</i> . . . . .	165
8.5. Exploration de la performance et de la consommation énergétique . . . . .	167
8.5.1. Mesure du système global (PDU) . . . . .	167
8.5.2. Mesure du système de stockage (capteurs) . . . . .	168
8.6. Conclusion . . . . .	169
<b>Chapitre 9. Performance et consommation des SSD : résultats expérimentaux</b> . . . . .	<b>171</b>
9.1. Introduction . . . . .	171
9.2. L'impact des E/S sur la performance et la consommation énergétique . . . . .	172
9.3. Une vue macroscopique de la performance et de la consommation des systèmes de stockage . . . . .	174
9.3.1. Configuration matérielle/logicielle pour les expérimentations . . . . .	174
9.3.2. Résultats des mesures . . . . .	175
9.3.2.1. Débit des HDD . . . . .	175
9.3.2.2. Débit des SSD . . . . .	177
9.3.2.3. Puissance d'un système avec HDD . . . . .	179
9.3.2.4. Puissance d'un système avec SSD . . . . .	181
9.3.2.5. Consommation énergétique d'un système avec HDD . . . . .	181
9.3.2.6. Consommation énergétique d'un système avec SSD . . . . .	182
9.3.2.7. Conclusion . . . . .	183
9.4. Vue microscopique de la performance et consommation des systèmes de stockage . . . . .	183

9.4.1. Utilisation de microbenchmarks . . . . .	183
9.4.1.1. Configuration matérielle/logicielle pour les expérimentations . . . . .	183
9.4.1.2. Résultats des mesures . . . . .	184
9.4.2. Utilisation d'applications vidéo . . . . .	189
9.4.2.1. Configuration pour les expérimentations . . . . .	189
9.4.2.2. Résultats avec le PRESET <i>ultrafast</i> . . . . .	190
9.4.2.3. Résultats avec le PRESET <i>slow</i> . . . . .	193
9.5. Conclusion . . . . .	194

## **Partie 4. Mémoires émergentes et conclusions . . . . . 197**

### **Chapitre 10. Mémoires non volatiles émergentes . . . . . 199**

10.1. Introduction . . . . .	199
10.2. Intégration des NVM . . . . .	201
10.2.1. Intégration comme mémoire secondaire . . . . .	201
10.2.2. Intégration comme mémoire principale . . . . .	201
10.2.3. Intégration dans les caches de processeur . . . . .	202
10.3. Mémoire à changement de phase : la PCM . . . . .	202
10.3.1. Concepts de base . . . . .	202
10.3.2. Intégration de la PCM . . . . .	203
10.3.2.1. La PCM comme mémoire principale . . . . .	204
10.3.2.2. La PCM comme mémoire secondaire . . . . .	204
10.3.3. Questions ouvertes . . . . .	205
10.4. Mémoire magnétorésistive : la MRAM . . . . .	206
10.4.1. Concepts de base . . . . .	206
10.4.2. Intégration de la MRAM . . . . .	208
10.4.2.1. MRAM comme cache . . . . .	208
10.4.2.2. MRAM comme mémoire principale . . . . .	209
10.4.3. Questions ouvertes . . . . .	209
10.5. Mémoire ferroélectrique : la FeRAM . . . . .	210
10.5.1. Concepts de base . . . . .	210
10.5.2. Intégration de la FeRAM . . . . .	211
10.5.2.1. FeRAM comme mémoire principale . . . . .	212
10.5.2.2. FeRAM comme stockage secondaire . . . . .	212
10.6. Mémoire résistive : la ReRAM . . . . .	213
10.6.1. Concepts de base . . . . .	213
10.6.2. Intégration de la ReRAM . . . . .	215
10.6.2.1. ReRAM comme cache . . . . .	215
10.6.2.2. ReRAM comme stockage secondaire . . . . .	215
10.6.3. Questions ouvertes . . . . .	216
10.7. Conclusion . . . . .	217

<b>Conclusion</b> .....	219
<b>Bibliographie</b> .....	223
<b>Index</b> .....	243