ANNEE 2011

THESE

présentée à

L'UNIVERSITE D'AIX-MARSEILLE 1

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE

Ecole doctorale :	Sciences pour l'Ingénieur Mécanique, Physique
Spécialité :	Micro et Nano Electronique

par

Florian MOLIERE

Ingénieur ISEN

FIABILITE DES TECHNOLOGIES CMOS FORTEMENT SUBMICRONIQUES (DSM) POUR DES APPLICATIONS AVIONIQUES, SPATIALES ET MILITAIRES

Soutenue le 25 Novembre 2011 devant la commission d'examen :

Président :	Jean-Luc AUTRAN	IM2NP
Directeur de Thèse :	Alain BRAVAIX	ISEN-IM2NP
Rapporteurs :	Nathalie MALBERT Alain SYLVESTRE	IMS G2Elab
Examinateurs :	Bruno FOUCHER Philippe PERDU (Co-encadrant)	EADS CNES
Invités :	Christian MOREAU Patrick HEINS	DGA AIRBUS



Institut Matériaux Microélectronique Nanosciences de Provence UMR 6242 CNRS, Universités Paul Cézanne, Provence et Sud Toulon-Var

Les DSM sont à l'électronicien ce que la prose est à Monsieur Jourdain

Molière

A ma conjointe Kasia,

REMERCIEMENTS

Le travail de thèse présenté dans ce mémoire a été réalisé à Suresnes (92) au centre de recherche EADS Innovations Works, en collaboration avec le laboratoire IM2NP (Institut Matériaux Microélectronique Nanosciences de Provence) et le CNES (Centre National d'Etudes Spatiales) dans le cadre d'une convention CIFRE (Convention Industrielle de Formation par la Recherche).

En préambule à ce mémoire, je souhaiterais adresser mes remerciements les plus sincères aux personnes qui m'ont apporté leur aide et qui ont ainsi à leur manière contribué à l'élaboration de ce manuscrit.

Je souhaite remercier Madame Nadine BUARD pour m'avoir accueilli dans le département «Systèmes Electroniques» d'EADS IW. Je tiens à témoigner toute ma reconnaissance à Madame Catherine MUNIER ainsi qu'à Monsieur Florent MILLER pour m'avoir successivement accueilli et fait confiance au sein de leur équipe «Fiabilité des Assemblages et Intégration en Electronique» et «Semi-conducteurs et Sureté de fonctionnement d'Equipements électroniques».

Je remercie vivement Monsieur Alain BRAVAIX, professeur à l'ISEN-Toulon, pour avoir dirigé et encadré ces travaux de thèse. Je connais Alain depuis les classes de Math Spé où déjà à l'époque, j'avais pu apprécier sa rigueur, son professionnalisme ainsi que son expertise dans les domaines de la physique des semi-conducteurs et des technologies submicroniques. Cela a été un grand plaisir de travailler ensemble.

Je remercie très sincèrement Monsieur Philippe PERDU, senior expert en Microélectronique au CNES pour avoir co-encadré ces travaux avec autant d'intérêt et de disponibilité. La clarté de ses analyses ainsi que la pertinence de ses conseils m'ont permis de progresser. Qu'il trouve ici le témoignage de ma profonde gratitude.

Je témoigne toute ma reconnaissance à Monsieur Bruno FOUCHER, senior expert en Microélectronique et grand gourou de la fiabilité de l'électronique chez EADS. Par sa pédagogie, ses compétences ainsi que sa gentillesse, il a grandement contribué à l'élaboration et au bon déroulement de ces travaux. Je tenais tout particulièrement à saluer la patience dont il a fait

preuve pour répondre à mes très nombreuses questions et interrogations en particulier celles posées le vendredi soir avant le week-end. Merci encore Bruno.

J'adresse également mes remerciements aux personnes qui ont accepté de participer à mon jury de thèse : à Madame Nathalie MALBERT, maître de conférence au laboratoire d'Intégration du Matériau au Système de Bordeaux (IMS) ainsi que Monsieur Alain SYLVESTRE, maître de conférence au laboratoire de génie électrique de Grenoble (G2Elab) qui en qualité de rapporteurs, ont jugé mon travail et porté une grande attention à la relecture de ce mémoire, ainsi qu'à Monsieur Jean-Luc AUTRAN, directeur adjoint de l'IM2NP pour m'avoir fait l'honneur de présider ce jury.

Je remercie Monsieur Christian MOREAU pour ses commentaires et suggestions mais également pour s'être déplacé jusqu'à Marseille pour assister à la soutenance.

Je souhaite tiens à remercier AIRBUS pour avoir soutenu ces travaux de thèse et tout particulièrement Monsieur Patrick HEINS qui a accepté de faire partie du jury et les membres de son équipe Jérôme TOULZE et Valérie FARRES pour les nombreux échanges que nous avons eu au long de ces années. Je remercie de la même manière Monsieur Philippe PONS. J'en profite également pour remercier Monsieur Jean-Marie CHOPIN pour avoir été l'instigateur de ses travaux au sein d'EADS.

Au cours de ces travaux, j'ai eu le plaisir de collaborer avec le laboratoire du Centre Intégré de Microélectronique de Provence Alpes Côte d'Azur (CIMPACA) ainsi que le laboratoire THALES. A ce titre, je tiens à remercier Messieurs Kevin SANCHEZ et Lionel FORLI ainsi que Mademoiselle Hélène CHAUVIN.

J'exprime mes remerciements à mes collaborateurs au sein d'EADS : Messieurs Renaud FOUGERAT, Gérard SALVATERRA, Pascal RETAILLEAU, Jérémy GALINAT et François ROSATI.

Un remerciement spécial pour les personnes qui ont successivement partagé mon bureau, à savoir, Thomas BOUTARIC, Katell MOREAU ainsi que Nicolas GUIBBAUD et qui m'ont écouté parler pendant 4 ans de la fiabilité des composants fortement submicroniques et cela sous sourcilier. Merci à eux.

Je tiens également à remercier mes collègues de travail à savoir (dans l'ordre alphabétique pour pas faire de jaloux), Thomas AUGUSTE, Antonin BOUGEROL, Bastien DUONG, Fatima IERVOLINO, Louis LEFUR, Jean-Baptiste LIBOT, Olivier MAIRE, Grégor MASSIOT, Guillaume PARENT, Vincent ROUET, la bretonne Estelle SCANFF (Breizh o Veva) et Cecile WEULERSSE. Je remercie spécialement Agnès CHAILLOT ainsi que Antoine RENAULT pour l'aide précieuse qu'ils m'ont apporté sur les simulations en éléments finis.

Je salue mes camarades thésards qui à leur tour soutiendront leur thèse de doctorat prochainement. Par ordre de soutenance (je prends les paris), Sébastien MORAND, Alioune CISSE, Benoit FLEURY, Sabrine HOUSSANY et Thomas SANTINI.

Je remercie fortement les stagiaires que j'ai eu à encadrer pour l'aide qu'ils ont pu m'apporter au cours de ces travaux, à savoir, Vardkès ASLANIAN et Glenn REBLEWSKI.

Je profite de cette occasion pour saluer les anciens collègues partis vers d'autres horizons, Romain BIENVENU, Nicolas FERRAND, Marc GRIEU, Arnaud DELYE, Jonathan LEROY, Mickael MANCEAU et Frédéric PORTA.

Enfin, je remercie chaleureusement ma conjointe Kasia ainsi que ma famille pour le soutien et la confiance qu'ils m'ont toujours accordés. Sans eux, cette étude n'aurait pas pu aboutir. Pour finir, j'ai une pensée émue pour mes amis d'enfance, en particulier pour le Dr. Jérémy VOLANTE qui a brillamment soutenu sa thèse de doctorat de médecine en Mai 2011 ainsi que pour mon excamarade de chambrée au Lycée Militaire d'Aix-en-Provence, Renaud BOUNICHOU qui malheureusement, nous a quitté trop tôt.

TABLE DES MATIERES

REMERCIEMENTS	5
TABLE DES MATIERES	9
LISTE DES ACRONYMES	
LISTE DES PARAMETRES	
INTRODUCTION GENERALE	

CHAPITRE	1 LES RISQUES INTRINSEQUES D'UN CIRCUIT INTEGRE FORTEMENT SUBM	ICRONIQUE 19
1.1 IN	TRODUCTION	21
1.2 LE	CIRCUIT INTEGRE	21
1.2.1	La connectique supra silicium BEOL	22
1.2.2	La zone active FEOL	23
1.3 IN	TEGRATION DES COMPOSANTS FORTEMENT SUBMICRONIQUES	24
1.3.1	Le nœud technologique d'un composant électronique	24
1.3.2	La filière technologique du composant	25
1.3.3	Les limites de l'intégration et de leur conséquences	
1.3.4	Bilan sur les conséquences et les limites de l'intégration	
1.4 IN	NOVATION TECHNOLOGIQUES	
1.4.1	Innovations BEOL	
1.4.2	Innovations technologiques FEOL	
1.4.3	Innovations technologiques architecturales	
1.4.4	Bilan des innovations technologiques	51
1.5 Co	DNCLUSIONS	53

CHAPITRE 2	LES MECANISMES DE DEFAILLANCE D'UN CIRCUIT INTEGRE FORTEMENT SU	IBMICRONIQUE 55
2.1 INTRO	ODUCTION	57
2.2 LES N	/IECANISMES DE DEFAILLANCE DU CIRCUIT INTEGRE	57
2.3 LES N	/ecanismes de defaillance BEOL	58
2.3.1	L'électromigration	58
2.3.2	La migration atomique induite par une contrainte hydrostatique	64
2.3.3	Le claquage d'oxyde intermétallique	68
2.4 LES N	/IECANISMES DE DEFAILLANCE FEOL	71
2.4.1	Le claquage de l'oxyde de grille	71
2.4.2	Les porteurs chauds	81
2.4.3	L'instabilité de la tension de seuil en température (BTI)	95
2.5 CON	CLUSION	

3.1	INTRODUCTION	
3.2	Profils de mission aeronautiques	
3.	2.1 Profil de mission d'un calculateur	
3.	2.2 Autres profils de mission	
3.	2.3 Spécifications en fiabilité	
3.3	APPLICATION DES MODELES DE DEGRADATION POUR TROIS APPLICATIONS AERONAUTIQUES	
3.	3.1 Les mécanismes de défaillance BEOL	
3.	3.2 Les mécanismes de défaillance FEOL	146
3.	3.3 Comparaison des critères de défaillances	164
3.4	Conclusion	

DEMONSTRATION, VALIDATION EXPERIMENTALE ET METHODOLOGIQUE POUR LA SELECTION **CHAPITRE 4** 4.1 4.2.1 4.2.2 Description de la technologie du composant testé176 4.2.3 Description des techniques de vieillissement......179 4.2.4 4.3.1 4.3.2 Application de la méthodologie sur deux mémoires flash NAND210 4.3.3

CONCLUSION GENERALE	217
PUBICATIONS ET ACTIVITES	223
REFERENCES BIBLIOGRAPHIQUES	225

LISTE DES ACRONYMES

Tormo	Description	Première	
Tenne	Description	occurrence	
AF	Acceleration Factor	Page 122	
AFM	Atomic Force Microscopy	Page 180	
AHI	Anode Hole Injection	Page 75	
ALD	Atomic Layer Deposition	Page 43	
AMD	Advanced Micro Devices	Page 42	
ASIC	Application Specific Integrated Circuit	Page 40	
BEOL	Back End Of the Line	Page 18	
BSIM	Berkeley Short-channel IGFET Model	Page 93	
CHC	Channel Hot Carrier	Page 57	
CIMPACA	Centre Intégré de Microélectronique de Provence Alpes Côte d'Azur	Page 180	
CMOS	Complementary Metal Oxide Semiconductor	Page 22	
CNES	Centre National d'Etudes Spatial	Page 183	
DIBL	Drain Induced Barrier Lowering	Page 30	
DPN	Doped Plasma Nitruration	Page 96	
DRAM	Dynamic Random Access Memory	Page 25	
DSM	Deep Sub Micron Components	Page 15	
EADS IW	European Aeronautic Defense and Space company Innovation Works	Page 170	
EEPROM	Electrically-erasable programmable read-only memory	Page 81	
EES	Electron Electron Scattering	Page 91	
ELECTRE	Elaborated Electronic Component Tester for R ^{adiation} _{eliability} Effects	Page 186	
EM	Electromigration	Page 16	
ЕОТ	Effective Oxide Thickness	Page 26	
ESD	Electrostatic Discharges	Page 31	
FCDC	Flight Control Data Concentrator	Page 115	
FDSOI	Fully Depleted Silicon On Insulator	Page 219	
FEG	Field Effect Gunn	Page 22	
FEOL	Front End Of the Line	Page 18	
FIB	Focused In Beam	Page 65	
FMGEC	Flight Management Guidance Envelope Computer	Page 115	
FN	Fowler-Nordheim	Page 33	
FPGA	Field Programmable Gate Array	Page 22	
GIDL	Gate Induced Drain Leakage	Page 30	
HBD	Hard BreakDown	Page 77	
HCI	High Carrier Injection	Page 18	
HK	High-K dielectric	Page 43	
HP	High Power devices	Page 25	
HTOL	High Temperature Operating Life test	Page 202	
IL	Interface Layer	Page 44	
ILD	Intra Layer Dielectric	Page 57	
IMD	Intra Metal Dielectric	Page 68	

ITRS	International Technology Roadmap for Semiconductor	Page 24
LDD	Light Doped Drain	Page 28
LEM	Lucky Electron Model	Page 83
LK	Low-K dielectric	Page 39
LP ou LOP	Low Power devices	Page 25
LSTP	Low Stanby Power devices	Page 26
LVSILC	Low Voltage Stress Induced Leakage Current	Page 33
MBU	Multiple Bit Upset	Page 119
MOSFET	Metal Oxide Semiconductor Field Effect Transistor	Page 81
MTTF	MeanTime To Failure	Page 62
MVE	Multiple Vibrational Excitation	Page 91
NBTI	Negative Bias Temperature Instabilities	Page 18
NMOS	n-doped Metal Oxide Semiconductor	Page 23
OTF	On The Fly	Page 99
PB	Progressive Breakdown	Page 77
PBTI	Positive Bias Temperature Instabilities	Page 57
PDSOI	Partially Depleted Silicon on Insulator	Page 219
PMOS	p-doped Metal Oxide Semiconductor	Page 23
PTM	Predictive technology Model	Page 93
R-D	Reaction-Diffusion Model	Page 95
RTN	Rapid Thermal Nitruration	Page 96
SBD	Soft BreakDown	Page 77
SCE	Short Channel Effects	Page 30
SEU	Single Event Upset	Page 119
SILC	Stress Induced Leakage Current	Page 33
SIP	System in Package	Page 16
SIV	Stress Induced Voiding	Page 57
SM	Stress Migration	Page 16
SMA	SubMiniature connector version A	Page 186
SOC	System on Chip	Page 16
SOI	Silicon On Insulator	Page 46
SON	Silicon On Nothing	Page 47
SOS	Silicon On Saphir	Page 47
SRAM	Static Random Access Memory	Page 18
SVE	Single Vibrational Excitation	Page 91
TDDB	Time Dependent Dielectric Breakdown	Page 34
TMBF	Temps minimum de Bon Fonctionnement	Page 200
TSMC	Taiwan Semiconductor Manufacturing Company	Page 52
TSOP	Thin Small Outline Package	Page 176
TTF	Time To Failure	Page 89
UMC	United Microelectronics Corporation	Page 52

LISTE DES PARAMETRES

Paramètre	Unité	Description
C _{ox}	С	Capacité d'oxyde
E _A	eV	Energie d'activation d'un mécanisme en température
E	V.cm ⁻¹	Champ électrique latéral
EOT	nm	Epaisseur Equivalente d'Oxyde de grille
E _{ox}	MV.cm ⁻¹	Champ électrique à travers l'oxyde de grille
f	Hz	Fréquence
I _{Dsat}	mA	Courant de drain en saturation
I _{GIDL}	μA.μm⁻¹	Courant de fuite drain-grille
I _{IUNC}	mA	Courant de fuite de jonction
I _{OFF}	nA.µm ⁻¹	Courant de fuite sous seuil
J	MAcm ⁻²	Densité de Courant dans les métallisations BEOL
J_{G}	A.cm ⁻²	Densité de courant à travers l'oxyde de grille
L_{EFF}	nm	Longueur de grille effective du transistor
L _G	nm	Longueur de grille du transistor
N _A	cm ⁻³	Concentration de dopage
N _{IT}	cm ⁻²	Densité d'état d'interface
N _{OT}	cm ⁻³	Densité de charges piégées dans l'oxyde
Т	°C ou °K	Température
TMBF	heures ou années	Temps minium de bon fonctionnement
T _{ox}	Nm	Epaisseur physique de l'oxyde de grille
TTF	heures ou années	Temps de défaillance
V _{DD}	V	Tension d'alimentation
\mathbf{V}_{DS}	V	Tension Drain Source
\mathbf{V}_{Dsat}	mV ou V	Tension de saturation de drain
V _{GS}	V	Tension Grille Source
\mathbf{V}_{TH}	mV ou V	Tension de seuil de basculement du transistor
W _G	nm	Largeur de grille du transistor
X	nm	Profondeur de jonction des zones épitaxiées
α	S.D	Facteur de sollicitation structure ON / Equipement ON
β	S.D	Pente de la distribution de Weibull
γ	cm.MV ⁻¹	Paramètre d'accélération en champ électrique
ε	S.D	Permittivité diélectrique
λ	FITs	Taux de défaillance
μ_{EEE}	$cm^2.V^{-1}.s^{-1}$	Mobilité effective des porteurs

LISTE DES CONSTANTES

Constante	Valeur	Unité	Description
A*	120	A.cm ⁻² .K ⁻²	Constante de Richardson
ε ₀	8,85. 10 ⁻¹²	kg ⁻¹ .m ⁻³ .A ² .s	Permittivité diélectrique du vide
k	8,62.10-5	eV.K ⁻¹	Constante de Boltzmann

INTRODUCTION GENERALE

Les équipements électroniques employés dans des applications aéronautiques, militaires et spatiales tels que les calculateurs de vol et les dispositifs de sécurité, intègrent de plus en plus de composants commerciaux fortement submicroniques (*Deep Sub Micron ou DSM*). Si, l'appellation DSM n'est à ce jour pas encore standardisée, il est toutefois admis que cette dénomination désigne les circuits intégrés dont la plus petite demi-distance métallique intra silicium est inférieure à 100 nm *(il sera question alors de nœud technologique inférieur à 100 nm)*. Au premier abord, cette définition ne semble concerner qu'une petite minorité de composants, mais en réalité, elle concerne une grande majorité des composants actifs produits en fonderie dans le monde depuis le début des années 2000 (microprocesseurs, mémoires, processeurs graphiques etc..).

L'attrait des concepteurs de l'industrie aéronautique pour ces technologies se justifie notamment par leur bonne intégration, leurs faibles consommations, les gains substantiels en performance en vitesse (fréquence de fonctionnement) ainsi que les capacités de stockage conséquentes que peuvent apporter ces composants en fonctionnement par rapport aux technologies plus anciennes. A titre d'exemple, avec l'apparition des mémoires flash, les capacités de stockage d'information ont été multipliées par 1000 en seulement 10 ans. De la même manière, les capacités de calcul des microprocesseurs ont été multipliées par un facteur 40 sur la même période.

Ces gains importants en performance ont été rendus possibles grâce au développement de deux approches complémentaires. La première porte le marché des semi conducteurs depuis 40 ans étant connu sous le nom de *More Moore*. Cette approche consiste à réduire tous les 2 à 3 ans, les paramètres géométriques et électriques du motif élémentaire implanté sur le silicium, par un facteur d'échelle estimé à 0,7. En 2010, la plus petite finesse de gravure disponible pour les microprocesseurs est de l'ordre de 40 nanomètres. Parallèlement, certains fabricants de circuits intégrés ont eu recours à des changements de matériaux afin d'améliorer les fréquences de fonctionnement mais également pour limiter les courants de fuite de leur composant en opération. Ces 10 dernières années ont vu donc le cuivre et les oxydes à faible constante diélectrique s'imposer dans les interconnexions intra silicium. De la même manière, les oxydes à haute permittivité et le silicium étiré ont permis respectivement de limiter les courants de fuite et d'augmenter le courant de fonctionnement.

La deuxième approche qualifiée de *More than Moore* est complémentaire à l'intégration dans le sens où elle s'attache à paralléliser les briques technologiques déjà existantes pour améliorer les fonctionnalités du composant. Comme illustration, afin de multiplier les capacités de stockage des mémoires flash, il existe les architectures de composant dans lesquelles plusieurs puces sont assemblées dans un même boitier (*SIP ou System in Package*). A l'inverse, les systèmes sur puces (*SOC ou System on chip*) permettent d'incorporer plusieurs fonctionnalités ou blocs technologiques sur une même puce en silicium.

Toutefois, si ces deux approches ont permis d'accroitre les fonctionnalités du circuit, elles ont rendu les puces plus sujettes aux défaillances. En effet, l'intégration massive de différentes fonctions et la superposition de niveaux métalliques qui en découle ont fait grandement croitre la densité de métallisation dans un circuit intégré. En conséquence, la structure devient plus sensible aux défaillances par électromigration (EM) et par contrainte thermomécanique (SM). De la même manière, il devient impossible pour des problèmes de stabilité en tension et de marge de bruit, de maintenir un facteur d'intégration constant pour tous les paramètres physiques et électriques du transistor. La première répercussion de cette intégration dite inhomogène, se traduit par l'augmentation des champs électriques dans la structure et par l'échauffement de la puce par effet joule, ce qui peut mener à terme aux claquages francs des différents oxydes de grilles. La seconde conséquence est spécifique à l'utilisation des technologies fortement submicroniques où à cette échelle, les courants de fuite traversant les oxydes de grilles par effet tunnel deviennent prépondérants en utilisation. Cette observation traduit le fait que la performance est souvent antagoniste des critères de fiabilité long terme

Néanmoins, en complément des performances intrinsèques, les concepteurs d'équipements aéronautiques exigent de leurs composants un niveau de qualité et de fiabilité accru. D'un point de vue général, la fiabilité du composant électronique se définit comme l'aptitude qu'à ce même composant à assurer sa fonction dans un milieu donné pendant un temps donné. La notion de contrainte environnementale prend tout son sens en aéronautique où les équipements électroniques embarqués sont soumis à des contraintes environnementales bien plus sévères que lors d'applications classiques dits « grand public ». Le composant assemblé sur carte doit pouvoir résister aux agressions électriques, radiatives, thermiques et électromagnétiques extérieures en plus des contraintes électriques imposées par le fonctionnement nominal. Ces agressions extérieures sont des facteurs d'accélérations du taux de panne du composant et de son usure. Parmi elles, la température apparait comme le facteur le plus critique au niveau du circuit intégré. Si bien qu'il existe plusieurs gammes de qualifications établies par le fabricant: la gamme

commerciale (0 /85°C), la gamme industrielle (-40 /100°C) représentative d'une application avionique civile et enfin la gamme militaire (-55 /125°C) spécifique aux applications missiles et spatiales.

De la même façon, les critères de fiabilité sur le long terme sont beaucoup plus exigeants. Par exemple, les spécifications en fiabilité pour des composants de calculateurs de commande de vol utilisés dans des applications avionique civiles sont de l'ordre de 10⁻⁷ défaillances par heure composant sur une durée de 30 ans à une température d'utilisation fixée à 55°C. A titre de comparaison, ces mêmes composants sont qualifiés chez les fabricants pour une durée de 10 ans à 125°C.

Comment est-il possible de s'assurer de la longévité d'un composant pour une application pour laquelle il n'a pas été spécifiquement conçu? Dans un premier temps, l'évaluation de la fiabilité des composants est fondée sur les résultats de tests de qualification opérés chez les fabricants ou l'utilisateur puis elle doit être infirmée ou confirmée ensuite par le retour opérationnel. Les différents retours d'expérience sur les technologies antérieures au nœud 0,5 µm employées pour les applications avioniques civiles telles que l'A320, ont démontré qu'en dehors des défauts de fabrication, les composants utilisés tiennent leur spécification sur au moins 20 ans.

Toutefois, du fait de l'obsolescence et de l'évolution rapide du marché, bon nombre de ces composants éprouvés d'ancienne génération ne sont aujourd'hui plus disponibles dans le commerce. Pourtant comme nous l'avons dit, pour des raisons de performance ou de disponibilité, les équipements n'ont d'autre choix que d'intégrer des composants de nouvelles générations dans leurs systèmes et architectures avec les risques de fiabilité qui en découlent. Cependant, on observe que la durée de vie des transistors fortement submicroniques à haute température, se réduit progressivement lorsque le nœud technologique diminue [BRA 09], conjointement à l'apparition de défaillances de façon progressives ou brutales [HUA 11]. Ceci rend impératif d'établir la fiabilité des technologies récentes et leurs différentes sensibilités au vieillissement (usure progressive) et à la défaillance (brutale ou prévisible) pour les applications avioniques à fortes contraintes environnementales.

L'objectif de ce travail a été dans un premier temps de vérifier la durabilité – ou la résistance au vieillissement - au niveau des composants fonctionnels de génération 90 et 65 nm tels qu'ils seront utilisés dans l'aéronautique dans les toutes prochaines années. Le second objectif a été de quantifier les risques en termes de marge de durée de vie et de proposer une

méthodologie de sélection composant basée sur leurs principaux critères de fiabilité pour les différentes applications aéronautiques.

La première partie définit les risques intrinsèques aux composants électroniques fortement submicroniques. Dans ce contexte, les choix technologiques ainsi que leurs conséquences sur le type de défaillance qui peuvent potentiellement se produire sur les technologies avancées, seront présentés et discutés. De plus, cette partie exposera les feuilles de route des fondeurs mondiaux, les principales innovations technologiques introduites au cours des 10 dernières années, mais également celles à venir pour les prochaines générations actuellement en développement.

La deuxième partie de ce manuscrit propose une synthèse des mécanismes de défaillance prépondérants dans les zones de métallisation (BEOL) et dans la partie active (FEOL) de la puce. Elle prend soin de détailler l'influence des nouveaux matériaux sur les modèles de dégradations qui seront utilisés dans les chapitres suivants.

La troisième partie présente les spécifications en durée de vie ainsi que les profils de mission de 3 applications aéronautiques. Pour les trois profils de missions étudiées, ce chapitre analyse la criticité des mécanismes de dégradation au niveau silicium en fonction de la génération technologique et des matériaux du composant. Ceci a pour objectif d'une part, d'identifier pour une application donnée, les technologies à risque susceptibles de ne pas respecter les spécifications en durée de vie des équipements et d'autre part de proposer des conditions de fonctionnement en température et en sollicitation, de façon à *fiabiliser* des composants qui ne le seraient pas d'ordinaire.

La quatrième partie vise à fournir aux utilisateurs, des techniques et des outils d'évaluation de la fiabilité d'un composant, non plus au niveau du silicium mais au niveau du produit. Pour cela, des techniques de vieillissements expérimentales NBTI et HCI ont été évaluées et mis en application sur une mémoire SRAM de génération 90 nm. Cette étude permet notamment de valider au niveau d'un produit, les estimations de durées de vie des matériaux du circuit intégré issues du chapitre 3. Enfin, cette partie propose également une méthodologie de sélection de composants fortement submicroniques en fonction de la technologie et du profil de mission du composant en application.

1 Les risques intrinseques d'un circuit integre fortement submicronique

1.1	INTRO	DUCTION	21
1.2	LE CIRO	CUIT INTEGRE	21
1.2	.1 LA (CONNECTIQUE SUPRA SILICIUM BEOL	22
1.2	.2 LA 2	ZONE ACTIVE FEOL	23
1.3	INTEG	RATION DES COMPOSANTS FORTEMENT SUBMICRONIQUES	24
1.3	.1 LE M	NŒUD TECHNOLOGIQUE D'UN COMPOSANT ELECTRONIQUE	24
1.3	.2 LA F	FILIERE TECHNOLOGIQUE DU COMPOSANT	25
1.3	.3 Les	LIMITES DE L'INTEGRATION ET DE LEUR CONSEQUENCES	26
1	3.3.1	Conséquence sur le champ électrique	28
1	3.3.2	Conséquences sur les courants de fuite	29
1	3.3.3	Conséquence sur la puissance dissipée	36
1.3	.4 Bil/	AN SUR LES CONSEQUENCES ET LES LIMITES DE L'INTEGRATION	37
1.4	INNOV	/ATION TECHNOLOGIQUES	38
1.4	.1 Inn	IOVATIONS BEOL	
1	4.1.1	Métallisations supra silicium	38
1	4.1.2	Oxydes à basse constante diélectrique et barrières de diffusion	41
1.4	.2 Inn	IOVATIONS TECHNOLOGIQUES FEOL	41
1	4.2.1	Oxydes de grille	41
1	4.2.2	Grilles métalliques	44
1	4.2.3	Substrat silicium sur isolant (SOI)	46
1	4.2.4	Silicium contraint	48
1.4	.3 Inn	IOVATIONS TECHNOLOGIQUES ARCHITECTURALES	50
1	4.3.1	Les composants systèmes sur puce (SOC)	50
1	4.3.2	Les composant systèmes dans un boîtier (SIP)	51
1.4	.4 Bil/	AN DES INNOVATIONS TECHNOLOGIQUES	51
1.5	CONCL	USIONS	53

1.1 Introduction

Ce chapitre pose les notions de base du composant électronique qui seront ensuite utilisées dans ce manuscrit. Cette partie à pour vocation de répondre à la question : qu'est ce qu'un composant électronique fortement submicronique ? Les notions de nœud et de filière technologique seront ainsi introduites. Les problématiques principales liées à l'intégration ainsi que les différents choix technologiques proposés pour les résoudre seront développés de façon synthétique pour souligner les paramètres importants impliqués dans la fiabilité du transistor à la puce.

1.2 Le circuit intégré

D'une manière générale, les composants électroniques commerciaux sont constitués de l'association d'une connectique leur permettant d'être reliés aux pistes d'un circuit imprimé, d'un boitier d'enrobage plastique ou céramique, et enfin d'un circuit intégré sur lequel est implantée la fonction électrique du composant (Figure 1.1).



Figure 1.1: Schéma d'un composant électronique traversant en coupe

Le circuit intégré est composée de deux niveaux distincts, la connectique supra silicium BEOL et la zone active FEOL :

- La connectique supra silicium, qualifiée BEOL pour Back End Of the Line, assure la liaison entre les différents composants de la zone active entre eux mais permet également de les relier aux fils de connexion externes. La façon dont sont interconnectés les éléments de la zone active détermine la fonction électrique du circuit.
- La zone active, aussi appelée FEOL pour Front End Of the Line, est constituée principalement de transistors Métal-Oxyde-Semi-conducteur MOS et de diodes ainsi que de résistances et de condensateurs selon l'application voulue. Ces éléments constituent les briques de base du circuit.

1.2.1 La connectique supra silicium BEOL

Les interconnexions sont constituées d'un réseau de lignes métalliques, s'organisant sur plusieurs niveaux séparés par des couches de diélectrique. Les connexions entre les lignes de deux niveaux différents sont assurées par des plots métalliques appelés vias. Une description générale de la zone BEOL est présentée Figure 1.2. L'architecture BEOL est standard quelle que soit la filière : elle est classiquement décomposée en interconnexions globales, intermédiaires et locales.

- Les interconnexions globales permettent de relier les niveaux intermédiaires aux fils de connexion externes. Ces connexions permettent en général de distribuer l'horloge et la polarisation aux différents blocs fonctionnels. Elles sont en général plus larges que les autres types d'interconnexion,
- Les interconnexions intermédiaires assurent la connexion avec les lignes du premier niveau local et permettent de connecter les différents blocs logiques entre eux. De leur agencement dépend la fonctionnalité électrique du circuit,
- Les interconnexions locales permettent de prendre les contacts sur la grille et les zones actives du transistor par l'intermédiaire de vias, le plus souvent en tungstène ou en polysilicium. Elles relient par conséquent les zones actives entre elles et permettent de réaliser des cellules logiques élémentaires telles qu'une structure CMOS ou un point mémoire.



Figure 1.2: Niveaux d'interconnexions d'un composant FPGA de génération 65 nm, microscope électronique à balayage FEG, électrons secondaires, x 7 000

1.2.2 La zone active FEOL

La partie active des structures numériques contient les zones de dopage de la puce. Le positionnement ainsi que l'agencement des jonctions p-n, permet de réaliser des transistors NMOS et PMOS. D'un point de vue électrique, le transistor est assimilable à un interrupteur commandé par la différence de potentiel grille source V_{GS} . Si cette tension est supérieure à la tension de seuil du transistor, alors un courant noté I_{DSAT} circule de part et d'autre des zones de diffusion (Figure 1.3).



Figure 1.3: Fonctionnement d'un transistor MOS, microscope électronique à balayage FEG, électrons secondaires, x 300 000

Les performances du transistor résident principalement dans la fréquence de basculement entre les positions passante et non passante. La fréquence est donnée par la relation :

$$f = \frac{I_{Dsat}}{C_{TOT} \cdot V_{DD}} \tag{1.1}$$

où, I_{Dsat} , C_{TOT} et V_{DD} représentent respectivement le courant de saturation, la capacité de la porte équivalente et la tension d'alimentation.

Le courant de saturation est directement lié aux performances maximum que peut délivrer une technologie à la tension d'alimentation V_{DD} . Il s'exprime en fonction des paramètres caractéristiques du transistor (canal N) :

$$I_{Dsat} = \frac{C'_{OX} \cdot \mu_{EFF} \cdot W_{EFF} \cdot (V_{GS} - V_{TH})^2}{2 \cdot L_{EFF}}$$
(1.2)

Avec C_{OX} la capacité d'oxyde de grille par unité de surface, μ_{EFF} la mobilité effective des porteurs, W_{EFF} et L_{EFF} respectivement la largeur et la longueur effective de la grille et enfin V_{TH} la tension de seuil du transistor pour V_{GS} = V_{DD}.

1.3 Intégration des composants fortement submicroniques

1.3.1 Le nœud technologique d'un composant électronique

Depuis 1970, le circuit intégré a été le thème de nombreux sujet de recherche dont l'aboutissement a toujours été l'intégration d'un maximum de fonction sur une même surface de puce. Si bien qu'au cours de ces 40 dernières années, les résultats de ces recherches ont pu alimenter le marché des semi-conducteurs en technologies nouvelles au rythme d'une génération technologique de composant tous les trois ans. Depuis les années 2000, cette tendance s'est même accélérée pour les composants fortement submicroniques où, une nouvelle génération technologique voit le jour tous les 2 ans.

Toutefois, comment définir un nœud technologique d'un circuit intégré? L'externalisation des chaînes de production dans les années 90 a conduit les acteurs de la microélectronique à standardiser leur appellation et à définir des objectifs communs de production. Dans cette optique, l'association des industries microélectroniques américaines (SIA) créa en 1992 une feuille de route nationale de technologies de semi-conducteurs (NTRS) qui deviendra en 1998 la feuille de route internationale des semi-conducteurs (ITRS). L'ITRS (International Technology Roadmap for Semiconductors) est un organisme résultant d'un consensus de fabricants de la micro-électronique qui définissent et anticipent les définitions des différents paramètres et procédés technologiques utilisés dans l'avenir. Les informations fournies dans les différentes rubriques de l'ITRS, sont en perpétuelle évolution et doivent être constamment adaptées pour les différents nœuds technologiques. Ainsi, les différentes rubriques sont remises à jour à chaque début d'année.

La génération technologique ou nœud technologique telle qu'elle est définie par l'ITRS, fait référence à la demi-distance minimale séparant deux plots métalliques sur la première couche locale d'un circuit intégré. Si cette demi-distance caractéristique, appelée aussi $Pitch_{1/2}$ est inférieure à 100 nanomètres, alors le composant sera considéré comme étant fortement submicronique (Figure 1.4). Cette valeur représente le pas lithographique utilisé par les fabricants

et traduit la densité de transistor sur une même surface de puce comme niveau d'intégration. La demi-distance intermétallique sert maintenant de référence à l'intégration des composants.



Figure 1.4: Composant SRAM de génération technologique 90 nm, microscope électronique à balayage FEG, électrons secondaires, x 45 000

1.3.2 La filière technologique du composant

Depuis 1997, il est nécessaire de prendre également en compte la notion de **filière technologique** pour une application considérée qui traduit l'optimisation des paramètres technologiques ainsi que certains paramètres physiques pour une dimension nominale du transistor. Si le nœud technologique fixe l'espacement interligne du premier niveau de métallisation, la filière technologique fixera certains paramètres principaux comme la tension de seuil (V_{TH}), le courant maximum à la tension V_{DD} (I_{Dsat}) le type de profil de dopage dans le canal du transistor. De ce fait, il est classique de distinguer les feuilles de route selon la filière *hante performance* (HP) ou *basse consommation* (LP ou LOP).

Les dispositifs *haute performance* sont des composants conçus pour favoriser la performance en vitesse de fonctionnement au détriment de l'autonomie (Tableau 1.1). Cette catégorie regroupe toutes les applications qui peuvent être alimentées par le secteur, principalement les processeurs de micro-ordinateurs et de serveurs.

A l'inverse des technologies *haute performance*, les dispositifs *basse consommation* privilégient l'autonomie. Ce critère est la priorité des secteurs des télécommunications et des systèmes embarqués. Les composants les plus représentés dans cette catégorie sont les mémoires SRAM et les SDRAM, qui s'alimentent avec des tensions toujours plus basses pour limiter les courants de fuite.

Depuis 2003, il est courant de trouver dans la littérature l'appellation *basse consommation en veille* (LSTP). Cette dernière catégorie est une optimisation de la filière de la basse consommation qui alimente le composant uniquement lorsqu'il est sollicité par l'utilisateur. Le reste du temps, le composant est en veille. Les composants de cette filière sont employés dans les équipements de télécommunication portatifs de nouvelle génération tels que les téléphones multimédia ou les ordinateurs de poche par exemple. Les règles de conception des transistors de cette famille de circuits intégrés sont plus relâchées par rapport aux autres filières, notamment en ce qui concerne la longueur de grille et l'épaisseur équivalente de l'oxyde de grille (Tableau 1.1). Mais cela, tout en ayant un espacement de plots de contact identique.

Technologie CMOS 65 nm		F	Performances	Intégration		
Filière	Application	Fréquence de basculement intrinsèque MOS f (GHz)	Consommation I _{DSAT} (μΑ/μm)	Courant de fuite I _{OFF} /W _G (µA/µm)	Longueur de grille physique L _G (nm)	Epaisseur d'oxyde équivalente EOT (A)
HP	Serveurs	1563	1121	0,34	25	18,4
LP	Portables	840	563	9.10-3	32	18,4
LSTP	Téléphonie	493	465	3.10-5	45	25,2

Tableau 1.1: Influence de la filière technologique sur les paramètres de performance et d'l'intégration [ITR 07]

Certains concepteurs s'accordent même à mêler plusieurs de ces filières technologiques au sein d'une même architecture de puce. C'est notamment le cas des composants FPGA sub 90 nanomètres où des transistors haute performance et basse consommation coexistent et sont adressés indépendamment selon le profil de l'application.

L'ensemble de ces considérations signifie qu'une bonne estimation de la fiabilité FEOL impose de connaître la filière du composant en complément de la génération technologique.

1.3.3 Les limites de l'intégration et de leur conséquences

Depuis 1970 et jusqu'en 1998, les différents paramètres caractéristiques du transistor tels que les dimensions, le dopage et la polarisation, ont été réduits de manière homothétique par un même facteur α ' pour garantir un champ électrique constant, i.e. $V/L_G = (V/\alpha') / (L_G/\alpha')$ (Figure 1.5). Ce facteur est estimé égal à 1,42 avec pour objectif de réduire la taille du transistor par 2 tous les 6 ans et ainsi soutenir la loi de Moore (Tableau 1.2).



Figure 1.5: Réduction homothétique des dimensions du transistor [TAU 98]

Sur cette période, la miniaturisation plus ou moins homogène du circuit intégré a permis de contrôler et de contenir les mécanismes de défaillance du silicium pour que ceux-ci n'impactent pas fortement la durée de vie du composant fonctionnel vis-à-vis de son application.

Dopage du substrat :

 $N_A \rightarrow N_A . \alpha'$

Toutefois, la réalisation de circuits intégrés sub CMOS 0,25 μ m haute performance (HP) et basse consommation (LP), s'est heurtée à plusieurs murs technologiques. Le plus critique demeure sans doute la réduction inhomogène de la tension d'alimentation par rapport aux épaisseurs équivalentes des oxydes (EOT) et la longueur de grille (L_G). Il apparaît en effet qu'au niveau du cœur logique, il n'est plus possible de diminuer la tension d'alimentation V_{DD} aussi vite que l'épaisseur d'oxyde de grille et la longueur de grille pour respecter des règles de stabilité en tension (V_{DD}/V_{TH} > 3) et de bruit. Ceci s'est traduit par l'utilisation de deux facteurs d'intégration découplés: le facteur K pour les dimensions et le facteur α pour la tension d'alimentation. En considérant les générations CMOS HP 0,25 μ m jusqu'au 45 nm, l'épaisseur électrique équivalente d'oxyde de grille a été réduite d'un facteur 5 alors que la tension de polarisation l'a été d'un facteur 1,7 sur la même période (Tableau 1.2).

		f		-		
Annéo	Génération	HP		LP		
Annee	technologique	V _{DD} (V)	EOT (nm)	V _{DD} (V)	EOT (nm)	
1989	1 µm	5	18	5	18	
1992	0 , 5 μm	5	12	5	12	
1995	0 , 35 μm	3,3	8	3,3	8	
1998	0 ,25 μm	2,3 à 1,7	5	2,3 à 1,7	5	
2001	0 ,13 μm	1,4	1,3 à 1,5	1,2	1,8 à 2,2	
2004	90 nm	1,3	1,2 à 1,4	1,1	1,4 à 1,8	
2007	65 nm	1,2	1,1 à 1,4	1,1	1,3 à 1,6	
2009	45 nm	1,1	1	1,1	1,2 à 1,4	

Tableau 1.2: Réduction non homothétique des dimensions du transistor [ITR 07]

En conséquence, cette intégration non homothétique a contribué à augmenter les champs électriques d'un facteur α ', tandis que la densité de puissance augmente du facteur α '². Par conséquent, les conditions de vieillissement deviennent intrinsèquement plus sévères par rapport aux conditions où le champ est constant.

1.3.3.1 Conséquence sur le champ électrique

Au niveau de la zone active FEOL, un champ électrique latéral E_{LAT} se manifeste dans la structure lorsque le transistor passe d'un état bloqué à un état passant ou saturé. Ce champ est exprimé par l'expression suivante [HU 85]:

$$E_{LAT} = \frac{V_D - V_{Dsat}}{l_0} \tag{1.3}$$

Ce champ s'exprime communément par le rapport entre la différence des tensions de drain appliqué V_D et en saturation V_{DSAT} , et l_0 une longueur caractéristique dépendant de l'épaisseur d'oxyde physique T_{OX} , de la profondeur de jonction X_J et de la longueur effective de grille L_{EFF} donnée par la relation [CHU 88]:

$$l_0 = 0,22 \cdot T_{OX}^{\frac{1}{3}} \cdot X_J^{\frac{1}{2}}$$
 si $L_{EFF} > 0,5 \,\mu m$ (1.4)

$$l_0 = 0,017 \cdot T_{OX}^{\frac{1}{8}} \cdot X_J^{\frac{1}{3}} \cdot L_{EFF}^{\frac{1}{5}} \qquad \text{si } L_{EFF} < 0,5 \,\mu\text{m}$$
(1.5)

La réduction homothétique des dimensions a permis de maintenir le champ électrique latéral à une valeur constante afin de ne pas dégrader les caractéristiques I_D-V_D du transistor en saturation ainsi que la mobilité des porteurs (ce qui caractérise le comportement dit canal long.). Cependant, avec l'intégration en dimension plus poussée accompagnée de l'implantation de zones peu dopées (LDD), le champ électrique latéral peine à être maintenu constant et augmente intrinsèquement à partir de la filière CMOS 65 nm haute performance. Cette augmentation du champ électrique peut conduire à une aggravation du mécanisme de porteurs chauds pour ces technologies.

Le champ électrique vertical E_{ox} , exprimé en MV.cm-1, apparaît à travers l'oxyde de grille et peut s'approximer comme étant le rapport entre la tension d'alimentation appliquée sur la grille (pour $V_{GS} = V_{DD}$) et l'épaisseur électrique équivalente de celui-ci (EOT).

$$E_{OX} = \frac{V_G - V_{FB} - \Phi_S}{EOT} \tag{1.6}$$

avec V_G la tension appliquée sur la grille, V_{FB} la tension de bande plates, Φ_S le potentiel de surface et EOT l'épaisseur électrique équivalente d'oxyde.

La réduction non homothétique de la tension d'alimentation a pour conséquence d'augmenter de manière significative le champ électrique au sein des zones actives de la puce. Au cours des 20 dernières années, le champ électrique vertical a été multiplié par un facteur 3 pour atteindre 10,5 MV.cm-1 en 2007 pour les technologies 65 nm CMOS HP à substrat massif, se rapprochant ainsi dangereusement du champ de claquage intrinsèque du SiO2, estimé à 11-15 MV.cm-1 selon la qualité de l'oxyde [BER 00] [RAN 06]. Au-delà de cette valeur critique de champ électrique, l'oxyde de grille perd ses propriétés isolantes, mais surtout la fonctionnalité électrique du composant est endommagée de manière progressive et irréversible (Figure 1.6). La défaillance totale et irrémédiable est qualifiée de claquage franc et est typique des technologies submicroniques antérieures à 1998.



Figure 1.6: Evolution du champ électrique à travers la zone active FEOL

1.3.3.2 Conséquences sur les courants de fuite

L'intégration a considérablement augmenté l'intensité des courants de fuite présents au sein de la brique FEOL. Bien que déjà présents dans les technologies plus anciennes et considérés autrefois comme négligeables, les courants de fuite ont un impact prépondérant sur la consommation et la fiabilité en dessous de la génération 0,13 µm. Les composants électroniques

des générations sub CMOS 0,25 μ m possèdent des oxydes de grilles d'épaisseurs physiques très fines, classiquement inférieures à 3 nm. Cette épaisseur caractéristique est une grandeur seuil à partir de laquelle les courants de fuite par effet tunnel deviennent très importants. De plus, la réduction non-homothétique de la tension d'alimentation par rapport aux différents paramètres physiques de l'intégration, a participé à l'aggravation des courants de fuite au sein de la structure. Il existe plusieurs courants de fuite : le courant de fuite sous seuil I_{OFF}, le courant de fuite grillecanal I_G (ou densité J_G), le courant de fuite grille-zone de diffusion I_{GIDL} (I_{GD}) et enfin le courant de fuite de jonction I_{IUN}. I_{OFF} demeure néanmoins le courant de fuite le plus prépondérant.

1.3.3.2.1 Courant de fuite sous seuil I_{OFF}

Le courant de fuite sous seuil appelé I_{OFF} , est une grandeur linéique exprimée en $\mu A/\mu m$. Ce courant de fuite traduit d'une part, le déplacement de porteurs de la source vers le drain du transistor (diffusion sous seuil) et d'autre part, de la grille aux oxydes très minces par effet tunnel (Figure 1.7).

Il se définit en absence de polarisation sur la grille comme $I_{OFF} = I_{DS}$ ($V_{GS} = 0$, $V_{DS} = V_{DD}$) à partir de la caractéristique I_{DS} - V_{GS} quand le transistor passe du régime de déplétion à l'inversion faible (faible V_{DS}) jusqu'en forte saturation ($V_{DS} = V_{DD}$).



Figure 1.7: Courant de fuite sous seuil

La diminution de la longueur de grille L_G a entrainé au niveau transistor, l'augmentation de l'effet canal court (SCE) et l'abaissement de la barrière de potentiel source-drain due à la tension de drain (DIBL) [**BRA 03a**]. Au niveau électrique, cela se traduit par l'abaissement de la tension de seuil V_{TH} et par la diminution de la pente sous seuil S qui conduit à l'augmentation du courant de fuite I_{OFF} au cours des différentes générations technologiques (Figure 1.8).



Figure 1.8: Evolution du courant de fuite sous seuil par nœud technologique et par filière [ITR 07]

De plus, les hautes températures ainsi que la tension de polarisation jouent un rôle accélérateur sur la dérive de la tension de seuil V_{TH} et sur le courant I_{OFF} [KNO 07] (Figure 1.9).



Figure 1.9: Représentation du courant sous seuil normalisé pour un transistor 65 nm, pour différentes filières et enfin pour différentes condition d'utilisation (T°C et V_{DD})

L'accroissement du courant I_{OFF} à travers les générations technologiques affecte la fonctionnalité et la consommation. Toutefois, l'intensité du courant I_{OFF} imposée par l'intégration n'est en général, sauf décharges électrostatiques (ESD) ou les phénomènes de perçage, que faiblement affectée par le vieillissement du transistor par exemple soumis à une contrainte porteurs chauds. Sa valeur est 3000 fois inférieure au courant de fonctionnement nominal I_{ON} pour un transistor non dégradé CMOS 65 nm HP (Tableau 1.2). Ce n'est par conséquent pas assez important en absence de commande, pour former un canal d'inversion, nécessaire à rendre le transistor passant.

1.3.3.2.2 Densité de courant de fuite J_G

La densité de courants de fuite grille-canal J_G est une grandeur surfacique exprimée en A/cm². J_G traduit le déplacement de charges électriques entre la grille et le canal avec ou sans commande sur la grille (Figure 1.10).



Figure 1.10: Densité de courant JG pour le NMOS

La densité de courant J_G est accentuée avec la réduction de l'épaisseur d'oxyde T_{OX} mais est en revanche atténuée avec la diminution de la surface d'oxyde A_{OX} [**PIC 06**]. Malgré tout, l'intégration entre les nœuds à multiplié par 10 000 la densité de courant de fuite provenant de la grille pour les applications HP (Figure 1.11) [**PET 06**].



Figure 1.11: Evolution de la densité des courants de fuite Jg en fonction de l'épaisseur d'oxyde pour la tension de grille en régime d'accumulation (transistor NMOS)

Lorsque le substrat d'un NMOS est en accumulation ($V_{GS} < 0$), la densité de courant J_G est due au passage des électrons de la bande de conduction de la grille vers la bande de conduction du substrat. En faible inversion ($0 < V_{GS} < V_{TH}$), le canal d'inversion n'est pas encore formé mais un courant devient mesurable pour les oxydes d'épaisseur inférieures à 3 nm. Il s'agit

de courants de fuite engendrés à faible tension de grille (LVSILC) **[PET 06]**. Lorsque le substrat NMOS est en inversion forte ($V_G > V_{TH}$), J_G est causé par le transfert des électrons de la bande de conduction du substrat vers la bande de conduction de la grille **[PET 06]**. Comme ces courants de fuite sont assistés par les pièges dans l'oxyde de grille et activés par le champ électrique, on les nomme de type *Stress Induced Leakage Current* (SILC) **[DEG 05]**.

Si pour les générations plus anciennes, le mécanisme principal de transport était sous injections de porteurs en régime Fowler-Nordheim (FN) assistées par des pièges neutres présents dans le volume de l'oxyde, le mécanisme d'injection pour les technologies inférieures au CMOS 0,13 µm est différent. La réduction des épaisseurs d'oxyde en dessous de 3 nm a favorisé les injections de porteurs sous seuil, par effet tunnel direct comme le montre la Figure 1.11.

1.3.3.2.3 Courant de fuite drain-grille I_{GIDL} (Gate induced drain leakage)

Le courant de fuite drain-grille I_{GIDL} est un mécanisme résultant du champ électrique local important présent dans la région de recouvrement drain-grille en régime accumulation [**BOU 97**] [**XU 98**]. Les conditions les plus critiques dans le NMOS sont $V_{GS} = 0$ V et $V_D = V_{DD}$ (Figure 1.12).



Figure 1.12: Courant de fuite IGIDL pour le transistor NMOS et dans le transistor signal de la porte inverseuse.

Le courant de fuite grille drain est modélisable par la relation suivante :

$$I_{GIDL} = A \cdot E_S \cdot e^{\frac{-B}{E_S}} \tag{1.7}$$

où A et B sont des paramètres dépendants des conditions d'utilisation et où E_s représente le champ électrique entre la grille et le drain **[LOP 04]**:

$$E_{S} = \frac{V_{DG} - \Phi_{S}}{\frac{\epsilon_{Si}}{\epsilon_{SiO2}} \cdot T_{OX}}$$
(1.8)

avec V_{DG} la tension drain-grille, Φ_s le potentiel de surface dépendant des paramètres technologiques, ε_{sI} et ε_{sIO2} les constantes de permittivité du silicium et de la silice (respectivement égales à 11,9 et 3,9) et T_{OX} l'épaisseur physique de l'oxyde de grille.

Le mécanisme GIDL est dû au champ électrique vertical présent dans la région de recouvrement grille drain. En régime d'accumulation, les électrons de la bande de valence de la grille peuvent traverser l'oxyde par effet tunnel jusqu'à la bande de conduction du drain. L'injection peut être accentuée par les pièges générés par porteurs chauds, prés du drain à l'interface [LOP 04]. Les électrons minoritaires sont alors collectés dans le drain et réexpédiés dans le substrat sous l'effet du champ latéral, créant ainsi le courant I_{GIDL}. Toujours en accumulation, les électrons réexpédiés peuvent se recombiner avec les trous majoritaires du substrat créant ainsi localement une zone de charge d'espace [BOU 97].

L'effet GIDL est accentué par plusieurs facteurs de l'intégration. L'injection de porteurs de la bande de valence de la grille à la bande de conduction du substrat par effet tunnel direct, est favorisée par la diminution de l'épaisseur de l'oxyde de grille T_{OX} et par l'augmentation de la concentration de dopants dans le substrat par rapport au dopage des jonctions L_{DD} [YUA 08]. Comme la densité de courant J_G , le courant I_{GIDL} aggrave la consommation statique du transistor ainsi que sa sensibilité au mécanisme de défaillance TDDB [POM 00].

1.3.3.2.4 Courant de jonction de fuite I_{JUNC}

Le courant de fuite de jonction I_{JUNC} représente les courants de fuite entre le drain et/ou la source et le substrat lorsque le transistor est en phase de commutation (Figure 1.13). Lors du basculement, la différence de potentiel imposée entre la source et le drain permet aux électrons issus du courant en saturation de surpasser la barrière de potentiel de la jonction p-n. En régime établi, le courant de fuite de jonction est faible (0,01 pA) et peut être considéré nul.



Figure 1.13: Courant de fuite IJUNC pour le NMOS

 $I_{\mbox{\scriptsize JUNC}}$ est généralement représenté à partir de l'équation de Schottky :

$$I_{JUNC} = I_{Dsat} \cdot \left(e^{\frac{q \cdot V_{SD}}{k \cdot T}} - 1 \right)$$
(1.9)

avec I_{DSAT} le courant de drain en saturation, q la charge élémentaire de l'électron (1,60.10⁻¹⁹ C), V_{SD} la tension appliquée entre la source et le drain (mV), k la constante de Boltzmann (8,62.10⁻⁵ J.K⁻¹) et T la température exprimée en kelvin (K).

Le courant de jonction est proportionnel aux surfaces des zones de dopage. Ainsi, la réduction des surfaces et l'augmentation de la valeur du dopage a eu un effet bénéfique sur la réduction des fuites de jonction. De ce fait, en absence d'agression électriques/électrostatiques extérieures (EOS/ESD), les générations fortement submicroniques seront en principes moins sujettes à cette catégorie de courant de fuite.

1.3.3.2.5 Discussion de l'influence des différents paramètres d'intégration sur les courants de fuite

L'impact de l'intégration des paramètres caractéristiques sur l'évolution des différents courants de fuite a été évalué de manière qualitative pour des transistors NMOS de génération CMOS 130-45 nm, au niveau de la conception du circuit et au niveau de la fabrication en fonderie (Tableau 1.3). Il en ressort que la longueur de grille et l'épaisseur d'oxyde de grille sont les paramètres les plus critiques vis-à-vis de l'augmentation des courants de fuite et que la largeur de grille et la profondeur de jonction sont des paramètres secondaires peu impactant.

Pour limiter l'augmentation des courants de fuites, la tension d'alimentation peut être réduite ou ajustée. De cette manière, un compromis entre les paramètres L_G, T_{OX} et V_{DD} permet de limiter les

effets néfastes de courants de fuite à travers les filières haute performance et basse consommation. Toutefois, la tendance générale liée aux nœuds technologiques fortement intégrés voit une augmentation inéluctable des courants de fuite.

Niveau	Paramètres	I _{OFF} (nA/μm)	J _G (A/cm ²)	I _{GIDL} (μΑ/μm)	I _{JUNC} (mA)
	Longueur de grille $L_G \downarrow$	$\uparrow \uparrow$	aucun	Ť	Aucun
Conception Circuit intégré	Largeur de grille $W_G \downarrow$	↓ faible	aucun	↓ faible	\downarrow (surface)
0	Alimentation $V_{DD}\downarrow$	$\downarrow\downarrow$	$\downarrow \downarrow$	$\downarrow\downarrow$	\rightarrow
	Epaisseur d'oxyde $T_{ox} \downarrow$	↑ faible	$\uparrow\uparrow$	$\uparrow \uparrow$	Aucun
Fonderie	Profondeur de jonction $X_J \downarrow$	\downarrow	↑ faible	aucun	\rightarrow
	Dopage du substrat $N_A \uparrow$	↓	aucun	↑	\downarrow

Tableau 1.3: Influence des différents paramètres de l'intégration sur les courants de fuites (La notation \uparrow traduit l'augmentation du courant de fuite alors que le symbole \downarrow exprime une diminution, la double flèche indique que le phénomène est plus marqué)

Les courants de fuites peuvent être le témoin de l'augmentation des pièges à l'interface ou proches de l'interface Si/SiO_2 ce qui permet d'utiliser ces courants comme précurseurs du claquage doux ou progressif de l'oxyde de grille. Ce phénomène d'usure se traduit par l'altération des paramètres électriques du composant sans pour autant que l'oxyde soit totalement défaillant. Cette défaillance est caractéristique des technologies submicroniques récentes (CMOS 90 nm et en dessous).

1.3.3.3 Conséquence sur la puissance dissipée

La puissance totale dissipée par un transistor se décompose comme la somme de la puissance dynamique dans l'état ON (partie active) et de la puissance dans l'état OFF (partie statique). La puissance statique totale P_{STATIQUE} , est composée de la somme des puissances statiques issues des différents courants de fuite décrit précédemment:

$$P_{STATIQUE} = P_{STAT_I_{OFF}} + P_{STAT_J_G} + P_{STAT_GIDL} + P_{STAT_J_{IUNC}}$$
(1.10)

La réduction non homogène de la tension de polarisation par rapport à T_{ox} et à L_{G} , a multiplié respectivement par 10 la puissance statique due au courant de fuite à travers la grille et par 15 la puissance statique due au courant sous seuil entre les générations 0,13 µm et 45 nm. L'impact des courants de fuite devient si prépondérant que la puissance statique de fuite P_{STATIQUE}
représente près de 40% de la puissance fournie pour le transistor CMOS 65 nm HP, alors qu'elle en représentait à peine 8% pour la génération 0,18 µm (Figure 1.14).



Figure 1.14: Répartition de la puissance fournie, composants CMOS, puce 10x10mm, 105°C [PER 06]

La puissance fournie au transistor ($P_{Fournie}$) se compose de la puissance dynamique à la commutation ($P_{Dynamique}$), de la puissance statique incluant les différents courants de fuite ($P_{Statique}$) et enfin de la puissance dissipée liée à l'effet Joule ($P_{Dissipée}$). La compensation de la puissance statique par une puissance fournie importante engendre par la même occasion l'augmentation de la puissance dissipée par effet Joule dans les interconnections BEOL toujours plus denses (quelques km sur une puce de microprocesseur 45 nm). La dissipation de puissance est surtout problématique pour les circuits haute performance tels que les microprocesseurs, les FPGA et même certaines mémoires où des solutions de refroidissement par air ou par eau doivent être envisagées.

1.3.4 Bilan sur les conséquences et les limites de l'intégration

La réduction non homothétique des paramètres classiques de l'intégration aggrave la consommation statique du circuit intégré. D'une part, le champ électrique augmente entre chaque nœud technologique, augmentant les risques de claquage d'oxyde de grille (FEOL) et d'oxyde intermétallique (BEOL) dans le silicium. D'autre part, les courants de fuite augmentent, impliquant une hausse conséquente de la consommation statique ainsi que de la puissance dissipée.

Le transistor MOS atteint ainsi ses limites en termes d'intégration et de consommation. Pour continuer la course à la performance tout en contenant les courants de fuite à un niveau acceptable, certains fabricants de composants ont eu recours à des changements de matériaux dans les zones FEOL et BEOL du circuit intégré. Ces innovations technologiques sont présentées dans la section suivante.

1.4 Innovation technologiques

1.4.1 Innovations BEOL

1.4.1.1 Métallisations supra silicium

1.4.1.1.1 Métallisation des circuits intégrés post 130 nm

Depuis l'avènement des circuits intégrés dans les années 70, les matériaux utilisés pour la réalisation d'interconnexion CMOS n'ont que peu changé. L'aluminium fait office de matériau conducteur alors que les vias reliant les différents niveaux sont réalisés en tungstène. Les barrières de diffusion sont élaborées à partir de l'association du titane Ti et de nitrure de titane TiN. Le dioxyde de silicium est lui utilisé en tant que diélectrique d'interligne. Ses propriétés d'isolant (de l'ordre de 10^{15} à $10^{16} \Omega$.cm), son gap important (8,9 eV), sa faible conductivité thermique (1,38 W.m⁻¹.K⁻¹) ainsi que sa faible permittivité ont fait de la silice l'oxyde historiquement utilisé pour isoler les lignes métalliques [MAN 98]. Une approximation ordinairement admise permet de modéliser le réseau comme un circuit électrique RC, R étant la résistance de la ligne métallique et C la capacité de couplage intermétallique.



Figure 1.15: Modélisation électrique d'une interconnexion

Dans cette configuration, le délai de propagation τ est proportionnel à la constante de temps RC du circuit définie par l'expression (1.11) :

$$\tau \propto R \cdot C \approx \in_0 \cdot \in_{0X} \cdot \rho \cdot \frac{L^2}{t \cdot d}$$
(1.11)

avec ε_0 la constante diélectrique du vide, ε_{OX} la permittivité de l'oxyde intermétallique, ϱ , L, t respectivement la résistivité, la longueur et l'épaisseur de la connexion métallique et d l'espacement intermétallique. Toutefois, l'accroissement continu de la densité d'interconnexion dans la zone BEOL additionnée à l'augmentation du nombre de niveaux de métallisation, a eu pour conséquence d'augmenter considérablement le temps de propagation du signal (Figure 1.16) **[ITR 07]**. Cette augmentation de délai est d'autant plus problématique qu'elle limite fortement la montée en fréquence du circuit. Ainsi, pour pouvoir contrebalancer cette limitation en performance, les fabricants de semi-conducteurs ont dû avoir recours à des remplacements de matériaux en utilisant le cuivre et les oxydes à basse permittivité en remplacement de l'aluminium pour les métallisations et du SiO₂ pour les isolations intermétalliques.



Figure 1.16: Délai de propagation BEOL et FEOL pour des interconnexions en aluminium (+ SiO₂, k = 4) et en cuivre (+ oxyde à basse permittivité LK, k = 2). Calculs réalisés en considérant des longueurs et des épaisseurs de lignes respectivement égales à 43 µm et 8 µm, pour les deux technologies **[NTR 97]**

1.4.1.1.2 Métallisation et barrières de diffusion des circuits intégrés fortement submicroniques

L'augmentation de la résistance de la métallisation R a contraint un grand nombre d'acteurs de la microélectronique à mener des recherches pour remplacer l'aluminium par un conducteur de plus faible résistivité **[TOR 95]**, **[GUT 95]**, **[MAR 07]**. La première approche a été d'introduire des grains de cuivre et/ou du silicium dans les métallisations en aluminium pour former des solutions solides AlCu ou AlSiCu **[BEI 98]**, **[PEN 01]**. Cependant, devant les problèmes d'inhomogénéité de la structure et d'électromigration, les fabricants de semiconducteurs ont plutôt opté pour la solution cuivre (damascène ou dual-damascène). Le procédé damascène provient d'une technique employée à Damas il y a 2 500 ans pour la décoration des épées. Le cuivre est déposé par plasma dans les tranchées préalablement creusés dans l'isolant **[GUI 05]**. Les microprocesseurs, les ASIC et les FPGA sont les premiers composants à avoir adopté les interconnexions en cuivre dès la génération technologique CMOS 0,13 µm. Pour les autres composants telles que les mémoires, le passage au cuivre dépend fortement des feuilles de route des fabricants.

Toutefois, le cuivre est un contaminant pour le silicium et diffuse beaucoup plus dans l'oxyde d'isolation que l'aluminium dans les mêmes conditions par exemple si l'on compare le coefficient de diffusion $D = 5.10^{-5} a 5.10^{-4} cm^2/s$ pour le cuivre contre $10^{-14} a 4.10^{-10} cm^2/s$ pour l'aluminium **[WON 07]**. De manière à empêcher toute contamination dans les oxydes intermétalliques et limiter l'effet des délaminations entre les différents matériaux, une double barrière de diffusion Ta/TaN est employée pour les technologies cuivre sub CMOS 0,13 µm. Cette double barrière est nécessaire parce que le nitrure de tantale (TaN) possède de meilleures propriétés de barrière de diffusion que le tantale (Figure 1.17) **[GUI 05]**. En revanche le tantale possède une meilleure adhérence avec le cuivre que le nitrure de tantale, ce qui permet d'améliorer les performances en électromigration **[GUI 05]**.



Figure 1.17: Interconnexion vue en coupe sub CMOS 0,13 µm

Enfin, la métallisation est couverte par une couche de recouvrement (*Capping Layer*) qui fait office de couche d'arrêt de gravure lors d'étapes ultérieures (Figure 1.17). Cette couche sert également de protection contre l'humidité introduite également lors des phases de gravure humide. Le type de couche de recouvrement varie selon les fondeurs : il est courant de rencontrer des couches de recouvrement en nitrure de silicium SiN_x , en $SiC_xN_yH_z$, en tantale, nitrure de tantale Ta/TaN ou encore en CoWP [HU 06]. L'interface métal/couche de recouvrement constitue le principal site de diffusion des atomes métalliques, par conséquent le choix des matériaux est essentiel et déterminera la fiabilité des interconnexions BEOL.

1.4.1.2 Oxydes à basse constante diélectrique et barrières de diffusion

Le couplage capacitif intermétallique C, participe également à l'augmentation du délai de propagation à travers les couches BEOL. Les recherches menées à la fin des années 90 ont conduit à remplacer, pour les microprocesseurs, dès la génération CMOS 0,13 μ m, l'oxyde SiO₂ par un oxyde à permittivité plus faible dit *low k*. Classiquement, le diélectrique retenu est du dioxyde de silicium dopé au carbone SiOC ou SiOCH **[CHE 04] [CHE 07]**. Le basculement vers les oxydes à basse permittivité s'est opéré simultanément avec le remplacement des interconnexions en cuivre. Depuis, l'intégration impose l'utilisation d'oxydes possédant une permittivité diélectrique de plus en plus basse. En 2010, la plus faible permittivité rapportée dans la littérature est de 2,4 **[MAT 10]**.

1.4.2 Innovations technologiques FEOL

1.4.2.1 Oxydes de grille

L'intégration a toujours favorisé le couplage capacitif entre la grille et le canal en réduisant l'épaisseur physique de l'oxyde T_{OX} . Cela eut pour conséquence d'améliorer la fréquence de basculement mais aussi d'augmenter considérablement les courants de fuite à travers l'oxyde de grille (J_G et I_{GIBL}). Pour contenir les courants de fuite, les fabricants ont eu recours à des modifications technologiques importantes sur les procédés d'élaboration des diélectriques de grille.

1.4.2.1.1 Oxydes nitrurés

La première innovation au niveau de l'oxyde de grille fut introduite dès les générations CMOS 0,13-0,18 µm avec l'apparition des oxydes nitrurés SiON. Les oxydes nitrurés avaient été mis initialement en place pour protéger l'oxyde de grille des PMOS des contaminations avec l'élément bore, lors du dopage-n de la grille à haute température **[LEE 08]**.

L'étape de nitruration peut être suivie d'une étape de recuit à l'ammoniac NH_3 (500°C) ou au protoxyde d'azote N_2O (900 à 1200°C). Le but étant de diffuser des espèces N dans le volume de l'oxyde afin de recouvrir les liaisons pendantes Si- et Si-O en liaison Si-N à l'interface Si-SiO₂ et de limiter par la même occasion les états d'interface et le risque de défaillance par porteurs chauds **[TAK 82], [HE 05]**. La nitruration permet également d'augmenter la constante diélectrique de l'oxyde et le champ de claquage intrinsèque de l'oxyde de l'ordre de 3 % selon le procédé. Par conséquent les oxydes nitrurés sont un peu plus robustes que les oxydes SiO_2 , vis-à-vis de la défaillance par claquage franc **[WU 03]**. De plus, l'épaisseur physique T_{OX} plus importante des oxydes nitrurés par rapport à leur homologues en SiO_2 (à même EOT), leur permet à contrainte équivalente, de délivrer un moindre courant de fuite à travers la grille (SILC et LVSILC) et donc d'être moins sensibles aux claquages doux **[HU 06]**.

En revanche les couches nitrurées ont un impact négatif sur la dégradation de la tension de seuil des transistors PMOS soumis aux contraintes de type NBTI. Lors de l'étape de recuit, les espèces azotées viennent endommager et casser les liaisons Si-O présentes dans le volume de l'oxyde. Les liaisons pendantes Si- ainsi créées vont se comporter comme des pièges à trous en régime d'inversion forte et détériorer la tension de seuil de basculement des PMOS **[HU 06]**.

Toutefois, les oxydes nitrurés deviennent difficilement intégrables en dessous du nœud 65 nm pour lequel leurs épaisseurs atteignent 3-4 couches atomiques (CMOS 65 nm HP) et les courants tunnels deviennent alors prépondérants [CHA 05a].

1.4.2.1.2 Oxydes à haute permittivité Hi-K

Certains fabricants de microprocesseurs tels qu'Intel **[INT 03]** ou IBM-AMD **[IBM 07]** ont opté pour des oxydes à haute permittivité dit *high-k* (Hi-K) pour remplacer les oxydes nitrurés. Le but étant toujours de diminuer l'épaisseur d'oxyde équivalente EOT tout en maintenant une épaisseur physique T_{ox} constante de manière à limiter les courants de fuite.

Toutefois, les oxydes Hi-K ont des paramètres de mailles incompatibles avec le silicium. De ce fait, ils ne sont en général pas déposés seuls. Une couche d'interface en SiO_2 ou SiON est nécessaire pour adapter les différents paramètres de maille (Figure 1.18).



Figure 1.18: Empilement Grille/Oxyde HK/couche d'interface/silicium

De ce fait, l'épaisseur d'oxyde équivalente Hi-K ($EOT_{OX_{HK}}$) peut être exprimée par la relation (1.12) :

$$EOT_{OX_HK} = EOT_{IL} + T_{OX_HK} \cdot \frac{\epsilon_{SiO_2}}{\epsilon_{OX_HK}}$$
(1.12)

où EOT_{IL} est l'épaisseur d'oxyde équivalente de la couche d'interface, ε_{SiO2} et $\varepsilon_{OX_{HK}}$ représentent les constantes diélectriques de la silice et de l'oxyde à haute permittivité.

L'utilisation des oxydes à base d'hafnium a permis de réduire les courants de fuite à travers l'oxyde d'un facteur 1000 pour un PMOS et d'un facteur 25 pour un NMOS **[SEM 08]**. Les propriétés diélectriques ainsi que la stabilité thermique font du SiO_2 un très bon isolant (Tableau 1.4) et son remplacement n'est pas chose aisée. L'oxyde à haute permittivité doit répondre un certain nombre de critères, à savoir **[MCP 02]**:

- Une structure amorphe pour un dépôt niveau par niveau (ALD),
- Un gap E_G important qui garantit la nature isolante de l'oxyde ($E_G = E_{Conduction} E_{Valence}$),
- Un taux de défauts d'interface N_{IT} faible ou similaire aux oxydes SiO₂ et SiON (≈10¹¹ cm⁻²)
 [MAN 98],
- Une imperméabilité aux contaminations au bore lors des étapes de dopage,
- Une stabilité thermique avec le substrat (Si, Si-Ge) et avec la grille (poly Si, métal).
- Un champ électrique de claquage intrinsèque E_{BD} important. La relation entre la permittivité k et le champ de claquage intrinsèque E_{BD} est donné par **[RAN 06]**:

$$E_{BD} = 35 \cdot k^{-0.54} \tag{1.13}$$

Une étude comparative permet de lister les caractéristiques principales des oxydes HK et de les comparer avec celles du SiO_2 (Tableau 1.4). La plupart des oxydes à forte constante diélectrique possèdent soit un gap faible, soit un champ de claquage faible. Seuls les diélectriques à base d'hafnium semblent présenter des caractéristiques thermiques et diélectriques équilibrées. Les choix technologiques retenus par les fabricants pour le CMOS 45 nm HP, sont soit du HfO₂ soit du HfSiON. Les formules chimiques exactes sont gardés confidentielles par les fabricants.

Oxydes HK	Structure cristalline	Gap (eV) [JIA 05]	Offset de la bande de conduction (eV)	Offset de la bande de valence (eV) [JIA 05]	Constante diélectrique e	Champ de claquage E _{BD} (MV/cm)
SiO ₂	Amorphe	8,9	3,5	4,4	3,9	11-15
Al_2O_3	Amorphe	5,1	2,8	1,8	9	11,2-13,8
ЦЮ	Cubique	5,7	1,5	3,4	25	6,7
	Tétragonal	5,7	1,5	3,4	25	3,9
Τ. Ο	Cubique	4,5	0,3	3,1	26	3,7
1 a ₂ O ₅	Tétragonal	4,5	0,3	3,1	26	2,1
ZrO ₂	Cubique	7,8	1,4	3,3	29	5,7
	Tétragonal	7,8	1,4	3,3	29	3,3
T:O	Tétragonal	3,5	1,2	1,2	80-95	2,5
1102	Trigonal	3,5	1,2	1,2	80-95	1,0

Tableau 1.4: Comparaison des caractéristiques des oxydes à forte permittivité avec celles du SiO2

[MCP 03] [RAN 06]

D'un point de vue général, le claquage franc est un cas extrême où la perte de fonctionnalité d'un composant est causée par une défaillance physique de l'oxyde. Or, l'apparition des courants de fuite dans les oxydes des technologies sub CMOS 130 nm, a rendu ce mécanisme moins critique en comparaison des mécanismes à dérives paramétriques tels que le NBTI. Les oxydes high-k possèdent une épaisseur d'oxyde physique plus importante que leur homologue en SiON. Par conséquent, le risque de défaillance par un claquage franc semble moindre. De plus, la couche d'interface va servir de tampon et prévenir le claquage franc de l'empilement tout entier **[CHO 08].**

En revanche, les oxydes à haute permittivité tels que le HfO2 ou le HfSiON comportent initialement une grande quantité de charges et de pièges neutres en volume **[YOU 06]**. Cette quantité est beaucoup plus importante que pour les technologies à base d'oxydes nitrurés. Ces charges peuvent être piégées puis dépiégées et participer aux courants de fuite grille-canal et par la même occasion perturber la stabilité de la tension de seuil (NBTI). Nous verrons par la suite comment ce dédoublement d'oxyde de grille modifie le processus de claquage de l'empilement.

1.4.2.2 Grilles métalliques

Les grilles des transistors fortement submicroniques sont réalisés en poly silicium. Pour les NMOS par exemple, la grille est fortement dopée par des matériaux de type V selon la classification périodique des éléments de Mendeleïev. Dans cette configuration, les électrons sont intrinsèquement les espèces majoritaires dans la grille. Lorsque le transistor est passant, en l'occurrence $V_{GS}=V_{DD}$, les électrons majoritaires sont attirés alors que les trous minoritaires sont repoussés à l'interface grille-oxyde (Figure 1.19).



Figure 1.19: Tension de poly déplétion et diagramme de bande

Il se créée ainsi une zone de déplétion d'épaisseur W_{POLY} qui peut être décrite par la relation :

$$W_{POLY} = \frac{(\epsilon_0 \cdot \epsilon_{0X})^2 \cdot E_{OX}}{2 \cdot q \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot N_{POLY}}$$
(1.14)

avec E_{OX}, le champ électrique vertical et N_{POLY} la concentration de dopage de la grille.

D'un point de vue électrique, une chute de potentiel V_{POLY} est observée dans l'électrode en poly silicium **[INT 03]**. La tension V_{POLY} n'est cependant jamais prise en compte. De manière générale, cette épaisseur de déplétion se comporte comme une barrière isolante d'épaisseur W_{POLY} qui doit être ajoutée à l'épaisseur d'oxyde pour le calcul de l'épaisseur électrique équivalente d'inversion EOT_{INV} **[FAN 98]** :

$$EOT_{INV} = EOT_{OX} + W_{POLY} \tag{1.15}$$

De ce fait, la déplétion de la grille en poly silicium rend difficile l'ajustement de la tension de seuil V_{TH} . Pour lutter contre ce phénomène, les fabricants ont du augmenter le dopage de la grille N_{POLY} entre chaque nœud technologique. Or, les grilles sont déjà fortement dopées (de l'ordre de 10^{21} atomes/cm³ pour le CMOS 65 nm) et il devient impossible de les doper au-delà de 5.10^{22} atomes/cm³ (densité atomique du silicium).

Devant ce mur technologique, certains fabricants ont remplacé les grilles en poly silicium par des grilles métalliques. Les grilles métalliques présentent aussi l'avantage de supprimer le risque de pénétration du bore dans l'oxyde de grille et présente une meilleure compatibilité avec l'interface des oxydes à haute permittivité **[BOH 07]**.

Le challenge actuel pour les fabricants est de choisir des grilles métalliques présentant les mêmes valeurs de travail de sortie que les grilles n_+ et p_+ en poly silicium. Il existe en 2011, deux procédés de dépôt de grille métalliques. Le premier connu sous le nom de *Metal Last* fut développé par Intel pour le nœud CMOS 45 nm HP [MIS 07]. Ce procédé consiste à implémenter la grille métallique après avoir préalablement déposé l'oxyde à haute permittivité. Cette technique est principalement adaptée aux filières hautes performances et tout particulièrement aux technologies Si-Ge [HOF 10].

Le deuxième procédé qualifié de *Metal First* et fut développé par Sematech et l'alliance IBM. Il consiste à déposer la grille métallique avant les étapes de recuit des zones épitaxiées source et drain. Ce procédé semble adapté aux filières basses consommations en veille ainsi qu'aux technologies SDRAM. Toutefois, ce procédé semble montrer ses limites pour les PMOS disposant une épaisseur d'oxyde équivalente inférieures à 8 nm du fait de la difficulté d'ajuster les travaux de sortie pour ce type de transistor **[HOF 10]**.

L'impact du procédé de fabrication des grilles métalliques sur les mécanismes de défaillance est encore méconnu sur de longues durées d'utilisation et ne sera pas étudié aux cours de ces travaux.

1.4.2.3 Substrat silicium sur isolant (SOI)

Les premiers substrats SOI sont apparus au début des années 70 dans des dispositifs utilisés pour des missions militaires et spatiales. L'attrait des équipementiers aéronautiques pour ces technologies se justifie notamment par la bonne tenue aux radiations que peuvent apporter ces technologies à haute altitude. Les structures SOI sont maintenant utilisés pour les applications radio fréquences (une pente sous seuil quasi-idéale de 60 mV/décade à la température ambiante) et pour les dispositifs CMOS à basse consommation et haute performance. Elles ont aussi montrés des capacités pour les applications de puissance car ces structures disposent d'un échauffement moindre à courant équivalent.

Les substrats SOI sont des substrats en silicium sur lesquels une couche d'isolant a été déposée. Cet isolant peut être aussi bien du saphir (SOS), de l'air (SON) ou de la silice (SiO_2) . Les premiers circuits à substrats SOI furent réalisés à partir de structures partiellement déplétées avec une épaisseur du corps (body) sensiblement égale à la profondeur de jonction des zones dopées. Les structures plus récentes sont dites entièrement déplétées avec une épaisseur du body proche de la profondeur de déplétion pour favoriser la mobilité des électrons (Figure 1.20).



Figure 1.20: Coupe de microprocesseur 90 nm en technologie SOI partiellement déplétée, microscope électronique à balayage FEG, électrons secondaires, x 120 000

Les substrats SOI présentent plusieurs avantages par rapport à leur équivalent à substrats massiques. Tout d'abord, la surface de contact entre les zones de diffusion et l'oxyde enterré SiO₂ permet de limiter fortement les capacités parasites de jonction puisque la constante diélectrique de la silice est plus faible que celle du silicium ($\varepsilon_{si} = 11,9$ contre $\varepsilon_{siO2} = 3,9$). De cette manière, seule la composante latérale du courant de fuite de jonction est à prendre en considération et la consommation statique s'en trouve fortement réduite. Ainsi la capacité totale du transistor est diminuée de près de 20 % et le courant de fonctionnement est augmenté d'environ de 30 % par rapport aux substrats massiques de même génération [SHA 99]. De plus, la limitation des courants de fuite issus des jonctions permet l'utilisation des circuits intégrés SOI pour des gammes de température plus importantes (applications en électronique de puissance).

Ensuite, la diminution des capacités de jonctions entraine la suppression du risque de déclenchement du thyristor PNPN parasite présent dans les structures CMOS et par la même occasion le risque de défaillance par latch-up. De plus, les architectures SOI sont moins sensibles aux radiations que leur homologue en substrat massique. Premièrement, la couche d'oxyde de silicium enterré peut filtrer une partie des particules radiatives selon leur type et leur longueur

d'onde. Enfin, l'épaisseur de la couche de silicium superficiel des technologies SOI (body) est moins épaisse que pour les technologies à substrat massique. Les particules radiatives disposent alors de moins d'espace pour générer des paires électrons-trous et par conséquent, le nombre de charges générées jusqu'au drain devient beaucoup plus faible.

Les transistors dissipent de la chaleur au cours des multiples commutations. Cette chaleur est généralement évacuée à travers le substrat vers un dissipateur thermique. Toutefois, la couche d'oxyde enterrée dissipe beaucoup moins bien la chaleur que le silicium (la conductivité thermique à 300K du SiO₂ est égale à 1,75 W.mK⁻¹ contre 61,9 W.mK⁻¹ pour le silicium) **[ARI 05]**. Il se produit alors un phénomène d'auto-échauffement entre le canal d'inversion et la couche d'isolant ce qui peut par conséquent accélérer les mécanismes de défaillance de l'oxyde et accentuer l'instabilité de la tension de seuil en température (NBTI) comme le phénomène de claquage d'oxyde de grille (TDDB). Deuxième point, les tensions de seuil des technologies SOI sont dynamiquement plus basses que les tensions de seuil des transistors sur substrat massique car la polarisation flottante du body ne permet plus d'ajuster la tension de seuil à l'application souhaitée.

1.4.2.4 Silicium contraint

L'introduction du silicium contraint en tant que matériau dans la zone BEOL permet d'obtenir à pas lithographique égal, un courant de fonctionnement plus important par rapport aux structures classiques. Le silicium contraint permet d'augmenter la mobilité des porteurs dans la couche d'inversion en modifiant le paramètre de maille des atomes du canal ou des zones de diffusion.

Pour les transistors NMOS, une contrainte en tension doit être exercée de manière longitudinale et transversale au niveau du canal de manière à favoriser la mobilité des électrons. Pour cela, il existe deux techniques. La première consiste à intégrer au sein du silicium des atomes de germanium caractérisés par un paramètre de maille plus important que celui du silicium, dans le but d'étirer la structure. De cette manière, la couche de silicium monocristallin qui servira de support au canal d'inversion, adaptera son paramètre de maille à celui de la couche SiGe ainsi créée. Cette solution est en phase de recherche et n'est toujours pas industrialisée à ce jour (2011). L'autre technique consiste à étirer les mailles du canal en recouvrant le transistor NMOS d'une couche de traction de type nitrure de silicium SI_XN_Y (Figure 1.21). Cette solution est en autre utilisée par AMD pour ses processeurs Athlon 64 3800+ **[AMD 11]**.



Figure 1.21: Couche de traction sur transistor NMOS [AMD 11]



Figure 1.22: Transistor PMOS sur silicium contraint [GHA 03]

En revanche, le canal des transistors PMOS est contraint en compression de manière à augmenter la mobilité des trous. Pour ce faire, les zones épitaxiées en poly silicium sont remplacées par du silicium dopé au germanium. De cette façon, les zones de diffusion sont dilatées de part et d'autre du transistor, ce qui permet au canal d'être compressé de manière longitudinale (Figure 1.22). Cette solution est appelée eSiGe (*embeded SiGe*) et est actuellement employée par Intel. Il existe un deuxième procédé (qui peut être complémentaire à la solution eSiGe), qui consiste à recouvrir le transistor PMOS d'une couche de compression de façon à comprimer le canal de conduction [AMD 11].

Le germanium a été choisi comme dopant pour sa structure cubique et ses paramètres de mailles proches de ceux du silicium (Tableau 1.5). Le germanium monocristallin comporte des caractéristiques en mobilité plus importante que le SiGe et pourrait être utilisé à la place du SiGe. Cependant, son paramètre de maille diffère encore trop de celui de l'oxyde de grille SiO₂ ce qui pourrait engendrer des pièges à l'interface et dans le pire des cas des dislocations. De plus, la température de fusion du germanium est trop basse (937°C), ce qui pourrait poser des problèmes lors des différentes phases de dopage ou de recuit.

Paramètres	Si	Si _{0,75} Ge _{0,.25}	Si _{0,50} Ge _{0,50}	Si _{0,25} Ge _{0,5}	Ge
Paramètre de maille (A)	5,4310	5,4825	5,5373	5,5960	5,6575
Mobilité des électrons (cm ² .V ⁻¹ .s ⁻¹)	1500	2100	7700	3300	3900
Mobilité des trous (cm ² .V ⁻¹ .s ⁻¹)	450	812,5	1175	1537,5	1900
Température de fusion (°C)	1415	1295,5	1176	1056,5	937

 Tableau 1.5: Paramètres atomiques et électriques du Si, SiGe et Ge [VIR 02]

Toutefois, l'inhomogénéité de l'interface entre l'oxyde de grille et le substrat contraint peut induire des états d'interface. Ces pièges d'interface participent directement au processus de dégradation du transistor par porteurs chauds ou par instabilité de la tension de seuil en température. Ce mécanisme est donc plus critique pour les substrats étirés que pour les substrats massiques [JEO 09].

1.4.3 Innovations technologiques architecturales

1.4.3.1 Les composants systèmes sur puce (SOC)

En parallèle des évolutions BEOL et FEOL décrites comme un approche *More Moore*, il est possible d'optimiser les performances d'un composant actif en travaillant sur l'architecture du composant. Cette tendance est qualifiée de *More Than Moore*.

Une première solution consiste à incorporer plusieurs fonctions (IP) sur un même circuit intégré. Cette solution existe depuis le début des années 70 avec l'avènement des premiers processeurs et est connu sous le nom de système sur puce (Figure 1.23). Les IPs des composants SOC sont parfois issus de procédés de fabrication différents. Pour illustration, les FPGA de génération 40 nm associent 2 générations technologiques différentes en considérant les entrées/sorties d'une part et la logique d'autre part. De même, la logique est elle-même déclinée en filière technologique haute performance et basse consommation. Soit 3 architectures de transistors différentes présentes sur un même circuit intégré, ce qui peut différentier grandement la sensibilité de chaque architecture aux mécanismes de défaillance pire-cas.



Figure 1.23: Processeur SOC

Les risques concernant les SOC sont similaires à ceux des circuits intégrés classiques. Néanmoins, une difficulté consiste à connaitre précisément les caractéristiques des IP implantés sur la puce (génération et nœud technologique, matériaux) pour pouvoir établir leur niveau de sensibilité aux différents modes de défaillance.

1.4.3.2 Les composant systèmes dans un boîtier (SIP)

Une autre option consiste à paralléliser les puces au sein d'un même boitier. Il est par exemple possible de fabriquer une mémoire flash de 16 Gb à l'aide de 4 puces de 4 Gb sans modifier le nœud technologique ni même la surface de puce (Figure 1.24). Cette solution plus récente est qualifiée de système dans le boîtier (SIP).



Figure 1.24: Mémoire flash NAND SIP de génération technologique 50 nm

En plus des mécanismes de défaillance du silicium liés à la technologie des puces parallélisées, les risques concernant les SIP peuvent se situer au niveau du boitier. Sous l'effet de contraintes thermomécaniques, des contraintes résiduelles préexistantes sur l'empilement puce(s)/raidisseur(s)/embase peuvent provoquer des délaminations.

La prédiction des risques exige une connaissance de la technologie du composant. A titre d'exemple, pour un profil de mission donné, les mécanismes de défaillances inhérents à une mémoire flash monopuce de 16 Gb gravée en 50 nm seront bien entendu différents de ceux d'une mémoire SIP 4x4 Gb en 90 nm.

1.4.4 Bilan des innovations technologiques

L'architecture et les matériaux du composant ainsi que la fiabilité qui en découle diffèrent selon la fonderie où les circuits intégrés ont été fabriqués. En conséquence, <u>il n'existe pas de</u> feuille de route standard technologie de composant pour le composant puisque celles-ci dépendent du nœud, de la filière technologique et également de la stratégie du fabricant et du fondeur. A l'aide d'études bibliographiques ainsi que de nombreuses analyses de construction réalisées chez EADS IW, il a été possible de représenter les innovations introduites par les principaux fondeurs mondiaux en fonction des nœuds et des filières *haute performance* et *basse consommation* (Tableau 1.6 et Tableau 1.7).

Techno CMOS µProcesseur HP et µcontrôleur					U	SA							Tai	wan				(Chin	e	E	urop	be .
		Intel				Al	Alliance IBM				UMC			Т	SM	С		SMIC			ST Micro		
		90 nm	65 nm	45 nm	32 nm	90 nm	65 nm	45 nm	32 nm	90 nm	65 nm	45 nm	90 nm	65 nm	45 nm	32 nm	22 nm	90 nm	65 nm	45 nm	90 nm	65 nm	45 nm
	SOI	Ν	N	N	N	0	0	0	N	N	0	?	Ο	0	0	0	0	Ν	Ν	?	N	O *	N
	Grille métal	N	N	0	0	N	N	N	0	Ν	N	N	Ν	Ν	Ν	N	0	N	N	N	N	Ν	N
FEOL	Diélectrique High k	N	N	0	0	N	Ν	Ν	0	N	N	Ν	N	Ν	Ν	N	0	N	N	Ν	N	Ν	N
	Silicium étiré	0	0	0	0	0	0	0	0	Ν	0	?	N	?	0	0	0	N	?	?	N	Ν	N
PEOI	Pistes Cuivre	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BEOL	Diélectrique Low k	0	0	0	0	0	0	0	0	0	0	0	?	0	0	0	0	0	0	0	0	0	0
Architecture	SOC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	SIP	N	N	N	N	N	N	N	N	N	N	N	N	Ν	N	N	?	N	N	N	N	N	N

Tableau 1.6: Feuille de route technologique « microprocesseur et microcontrôleur » des principaux fondeurs mondiaux (O=OUI, N=NON, * sur demande) au 01/01/2011

			USA										Taiwan									South Korea				
Techno CMOS Memoires		C	Cypress			Micron				Freescale			UMC			TSMC					Samsung					
		130 nm	90 nm	65 nm	130 nm	90 nm	65 nm	50 nm	34 nm	130 nm	90 nm	65 nm	90 nm	65 nm	45 nm	90 nm	65 nm	45 nm	32 nm	22 nm	130 nm	90 nm	65 nm	50 nm	34 nm	
	SOI	Ν	Ν	Ν	Ν	Ν	Ν	Ν	Ν	Ν	Ν	N	Ν	Ν	?	Ν	Ν	Ν	O *	O *	Ν	Ν	Ν	Ν	Ν	
	Grille métal	Ν	N	N	N	N	N	N	N	N	N	N	N	N	?	N	N	N	N	0	N	N	N	N	N	
FEOL	Diélectrique High k	N	N	N	N	N	N	N	N	N	N	N	N	N	?	N	N	N	N	0	N	N	N	N	N	
	Silicium étiré	N	N	N	N	N	?	?	?	N	N	N	N	0	?	N	?	0	0	0	Ν	N	N	N	N	
PEOI	Pistes Cuivre	N	O *	0	O*	O*	0	O *	0	0	0	0	0	0	0	0	0	0	0	0	Ν	N	N	N	0	
BEUL	Diélectrique Low k	N	O *	0	O*	O*	0	O *	0	0	0	0	0	0	0	0	0	0	0	0	Ν	N	N	N	0	
Architecture	SOC	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
	SIP	N	N	N	N	N	0*	O *	O *	N	N	N	N	N	N	N	N	N	N	N	N	N	O*	0*	O*	

Tableau 1.7: Feuille de route technologique « mémoires » des principaux fondeurs mondiaux (O=OUI, N=NON, * sur demande) au 01/01/2011

L'analyse de ces tableaux permet de souligner quelques constatations importantes pour notre étude. Définir un composant par son nœud technologique est largement insuffisant. La connaissance de la filière et du fondeur est indispensable. De ce fait, les tableaux permettent de noter les points suivants :

Tout d'abord, le **nœud technologique** permet de déterminer chez un fondeur, la date d'introduction d'une innovation technologique. Par exemple, le silicium étiré n'a été introduit par Intel qu'à partir du nœud 65 nm.

Deuxièmement, un fondeur n'introduira pas les mêmes innovations pour ses filières technologiques dites à haute performance et à basse consommation. A titre d'exemple, pour les générations 90 à 45 nm, le fondeur TSMC utilise des substrats SOI uniquement pour sa filière haute performance. La filière basse consommation est, quant à elle, fabriquée sur substrat de silicium massique.

Enfin, les innovations ne sont pas toutes introduites lors de l'optimisation d'un même nœud technologique chez **les fondeurs**. Autrement dit, un processeur Penyn 45 nm de chez Intel ne sera pas équivalent « technologiquement parlant » à un processeur Phenom II d'IBM/AMD gravé en technologie 45 nm. Le premier utilise en effet des oxydes de grille à haute permittivité alors que le second a choisi de continuer à produire des composants à oxydes de grille nitrurés SION.

1.5 Conclusions

Les défaillances engendrées par la miniaturisation du motif élémentaire, peuvent être identifiées au niveau FEOL selon les générations technologiques des composants :

Les technologies anciennes (CMOS > $0,25 \mu m$) peuvent conduire à des défaillances franches de l'oxyde de grille, dues aux champs électriques importants présents à travers la structure. La défaillance électrique engendrée intervient à la suite de la défaillance physique de l'un des oxydes de la structure. Ces défaillances peuvent être mises en évidence lors des tests de qualification mais ne semblent plus être prépondérantes pour les technologies CMOS récentes.

Les nouvelles technologies submicroniques (< CMOS 130 nm) s'avèrent être en effet plus sensibles aux défaillances progressives et irréversibles des transistors. Ces nouveaux mécanismes induisent une altération de la fonctionnalité et de la consommation statique et dynamique du composant, qui apparaît préalablement à la défaillance physique de la structure (cas extrême). Ainsi, seul le niveau de la dérive des paramètres de fonctionnement (temps d'accès, tension de seuil, etc...) peut permettre d'évaluer et de fixer le critère de défaillance. Les tests de qualification classiques en tension et en température ne semblent plus être adaptés pour repérer ce nouveau genre de défaillance.

Au niveau des métallisations BEOL, la réduction inhomogène des dimensions du motif par rapport à la polarisation entraine entre chaque nœud technologique, une augmentation du champ électrique entre les interconnexions métalliques. A ce niveau, le champ électrique devient le facteur prépondérant (les oxydes ne sont pas suffisamment fins pour être soumis aux courants statiques de fuite). De plus, les oxydes à basse permittivité récemment introduits possèdent un champ de claquage intrinsèque ainsi que des propriétés mécaniques bien plus faibles que ceux du SiO₂ classiquement utilisé. Cela a contribué à l'émergence de mécanismes nouveaux autrefois insignifiants d'un point de vue fiabilité, comme par exemple le claquage de l'oxyde intermétallique.

Enfin, les limites de l'intégration du circuit intégré ont conduit les fabricants à proposer de nouvelles solutions technologiques au cours de ces 10 dernières années. Ces solutions technologiques sont optimisées pour la performance ou la faible consommation. Elles répondent principalement aux besoins des marchés de production de masse que représentent aujourd'hui la micro-informatique, l'électronique grand public et les télécommunications. Il existe aujourd'hui sur le marché, une multitude de composants fortement submicroniques caractéristiques d'une ou plusieurs innovations technologiques. Ces composants sont identifiables par leur **génération**, leur **filière technologique**. Mais, il est également nécessaire de considérer la **fonderie** d'où ces composants sont issus.

Chacune de ces innovations a modifié l'impact des mécanismes de défaillance dans les zones BEOL et FEOL du circuit intégré. Si bien que la fiabilité intrinsèque du composant s'en trouve modifiée de manière positive ou négative. De ce fait, l'utilisation de ces technologies sur le long terme soulève pas mal d'interrogations. La grande diversité des technologies de composants disponibles sur le marché, impose au concepteur d'équipement aéronautique de connaitre les modes et les mécanismes de défaillance mais également les risques en fiabilité propres à ces technologies vis-à-vis de son application. Dans cette optique, le chapitre 2 de ce manuscrit décrit et évalue les mécanismes de défaillance BEOL et FEOL à travers 3 profils de mission caractéristiques d'applications aéronautiques.

2 Les mecanismes de defaillance d'un circuit integre fortement submicronique

2.1	INTROE	DUCTION	57
2.2	LES ME	CANISMES DE DEFAILLANCE DU CIRCUIT INTEGRE	57
2.3	LES ME	CANISMES DE DEFAILLANCE BEOL	58
2.	3.1 L'EL	ECTROMIGRATION	58
	2.3.1.1	Physique de la défaillance	59
	2.3.1.2	Longueur de Blech	60
	2.3.1.3	Impact du choix du matériau et de la génération technologique	61
	2.3.1.4	Modèle de dégradation	62
2.	3.2 LAN	/IGRATION ATOMIQUE INDUITE PAR UNE CONTRAINTE HYDROSTATIQUE	64
	2.3.2.1	Physique de la défaillance	64
	2.3.2.2	Modèle de vieillissement associé	65
2.	3.3 Le c	LAQUAGE D'OXYDE INTERMETALLIQUE	68
	2.3.3.1	Physique de la défaillance	68
	2.3.3.2	Modèle de dégradation	70
	2.3.3.3	Oxydes à basse permittivité	71
2.4	LES ME	CANISMES DE DEFAILLANCE FEOL	71
2.	4.1 LE C	LAQUAGE DE L'OXYDE DE GRILLE	71
	2.4.1.1	Oxydes SiO2 et SiON	72
	2.4.1.2	Oxydes à haute permittivité HK	76
2.	4.2 LES	PORTEURS CHAUDS	81
	2.4.2.1	Physique de la défaillance pour les oxydes SIO2, SION et les oxydes à haute permittivité	82
	2.4.2.2	Modèle de défaillance associé au porteurs chauds	88
2.	4.3 L'IN	STABILITE DE LA TENSION DE SEUIL EN TEMPERATURE (BTI)	95
	2.4.3.1	Physique de la défaillance pour les oxydes en silice et nitruré	97
	2.4.3.2	Modèle de prédiction de technologie (PTM BTI)	102
25	CONCU	ISION	107
	CONCE		±0/

2.1 Introduction

La miniaturisation des motifs élémentaires ainsi que l'introduction de nouveaux matériaux dans la fabrication des circuits intégrés, ont conduit les fabricants à devoir modéliser et anticiper les principaux modes de défaillances des circuits. La défaillance d'un dispositif submicronique ou d'un bloc fonctionnel n'est plus uniquement liée à une défaillance physique et franche mais également à une dérive des paramètres de fonctionnement au cours du temps. L'amplitude de ces dérives est généralement liée à la qualité de la technologie développée pour de tels composants même si ces phénomènes d'usure (wearout) sont directement dépendants des mécanismes physiques impliqués. Ainsi ces deux thématiques sont intimement liées, ce qui nécessite une approche détaillée des principaux paramètres représentatifs des différents modes de défaillance.

Le chapitre 1 a montré la grande diversité des technologies utilisées pour les composants disponibles sur le marché. Le chapitre 2 est consacré à la description des mécanismes physiques inhérents aux défaillances BEOL et FEOL du circuit intégré que nous utiliserons par la suite dans le chapitre 3 pour distinguer l'importance des différentes défaillances pour des applications avioniques et aérospatiales.

2.2 Les mécanismes de défaillance du circuit intégré

Les principaux mécanismes potentiels de défaillance dans un circuit intégré sont au nombre de six. Il est possible de distinguer sur la Figure 2.1 *(i)* les mécanismes opérants au niveau des lignes de connexion intra silicium (BEOL) comme l'électromigration (EM), le claquage des oxydes Inter-métal (TDDB, ILD), des contraintes en Stress Induced Voiding (SIV), puis *(ii)* les mécanismes liés au fonctionnement des transistors (FEOL) comme le claquage du diélectrique de grille (TDDB), les injections de porteurs chauds (CHC) et enfin l'instabilité en tensions de seuil (NBTI, PBTI).



Figure 2.1: Sites de défaillances potentiels dans un circuit intégré

2.3 Les mécanismes de défaillance BEOL

2.3.1 L'électromigration

L'électromigration est un mécanisme d'usure intervenant dans les connexions intégrées de la zone BEOL. Ce mécanisme traduit la migration d'atomes des grains dans un conducteur métallique qui apparait avec l'augmentation de la densité de courant maximum dans les lignes (Figure 2.2). Le déplacement des atomes engendre alors des lacunes, formant des cavités à la cathode ainsi qu'une accumulation de grain à l'anode qui peuvent migrer au niveau des barrières de diffusion et des vias **[LIN 07]**. Les lacunes se manifestent par une augmentation de la résistance du matériau (jusqu'au circuit ouvert) alors que les accumulations de matière peuvent créer des courts-circuits entre deux lignes voisines. Ce dernier cas demeure toutefois improbable pour les interconnexions de circuit intégré, les barrières de diffusions en bicouche de nitrure de tantale (TaN/Ta) protègent le diélectrique interligne de toute contamination provenant des lignes métalliques.



Figure 2.2: Mécanisme d'électromigration dans un conducteur

2.3.1.1 Physique de la défaillance

En l'absence de polarisation, les atomes de la piste métallique ne sont soumis à aucune force et la valeur du flux est considérée comme nulle. En revanche, lorsque l'interconnexion est sollicitée électriquement, deux forces intrinsèques viennent alors s'exercer sur les atomes de la structure:

Une force extérieure F_{ELECTRIQUE} du flux d'électrons sur les ions positifs provoquée par la circulation du courant électrique. Cette force est exprimée par la relation (2.1):

$$\vec{F}_{ELECTRIQUE} = e \cdot Z_{ION} \cdot \vec{E} = e \cdot Z_{ION} \cdot \rho \cdot \vec{j}$$
(2.1)

où e représente la charge élémentaire, Z_{Ion} et ϱ le numéro atomique des ions et la résistivité linéique du conducteur métallique (Ω .cm), j la densité du courant (A.cm⁻²).

Une force de friction mécanique F_{MECANIQUE} induite par le gradient de contrainte entre la ligne métallique et les barrières de diffusion/passivation d'une part et les vias d'autre part. Cette force est traduite par l'expression (2.2):

$$\vec{F}_{MECANIQUE} = -\Omega \cdot \frac{\partial \sigma}{\partial L}$$
(2.2)

où Ω est le volume d'atome déplacé (cm³), σ la contrainte mécanique (Pa) et L la longueur de la ligne (cm).

L'électromigration se traduit par le déséquilibre entre la force de friction dominante et la force électrostatique. Par conséquent la force résultante $F_{TOTALE} = F_{elec} + F_{friction}$ est orientée suivant le déplacement des électrons. Dans cette configuration, la vitesse d'entrainement de matière dans le volume v_{EM} est exprimée par l'expression (2.3) [GIG 03]:

$$V_{EM} = \frac{D_{EFF}}{k \cdot T} \cdot F_{TOTALE} \approx \frac{D_{EFF}}{k \cdot T} \cdot (e \cdot Z^*) \cdot \rho \cdot j$$
(2.3)

où D_{EFF} le coefficient de diffusion des espèces (cm².s⁻¹) en assimilant les atomes à des ions métalliques, kT l'énergie thermique, Z^{*} la charge effective (Z^{*} = Z_{Ion}- n_e $\sigma_e \lambda_e$) où n_e est le nombre d'électrons en collision avec les atomes, de conductivité σ_e et libre parcours moyen λ_e .

La valeur du coefficient D_{EFF} diffère selon le chemin local de diffusion. De cette manière, la diffusion effective D_{EFF} se compose d'une somme des diffusions élémentaires dans le volume (D_{VOL}) , à la frontière des grains (D_{GRAIN}) et aux différentes interfaces (D_{INT}) :

$$D_{EFF} = \sum_{i} D_i = D_{VOL} + D_{GRAIN} + D_{INT}$$
(2.4)

Les diffusions élémentaires dépendent de la nature du matériau (aluminium ou cuivre) et sont accélérées en température de manière indépendante selon la loi d'Arrhenius :

$$D_i = D_0 \cdot e^{\frac{-E_A}{k \cdot T}} \tag{2.5}$$

où D_0 est la constante de diffusion dépendante du matériau (cm².s⁻¹) et E_A l'énergie d'activation de la diffusion selon le chemin considéré (eV).

2.3.1.2 Longueur de Blech

Expérimentalement, il existe une longueur de ligne limite, appelée longueur de Blech, en deçà de laquelle le risque de défaillance par électromigration est nul **[BAO 03]**. Pour cette longueur de ligne, la force mécanique égale la force électrique. Ce qui implique que la vitesse de dérive de grain devient nulle :

$$V_{EM} = \frac{D_{EFF}}{k \cdot T} \cdot \left[(e \cdot Z^*) \cdot \rho \cdot j - \Omega \cdot \frac{\partial \sigma}{\partial L} \right] = 0$$
(2.6)

Dans ces conditions, il devient possible d'extraire le produit densité de courant critique-longueur de ligne critique qui est une caractéristique du nœud technologique. Cette relation est exprimée par la relation (2.7) :

$$(j \cdot L)_{CRITIQUE} = \frac{\Delta \sigma \cdot \Omega}{e \cdot Z^* \cdot \rho}$$
(2.7)

Ce produit est estimé à 1260 A/cm pour l'aluminium [BLE 76] et de 3700 A/cm pour le cuivre [OGA 01]. De ce fait, la connaissance de la valeur de la densité de courant présente dans l'interconnexion permet d'extraire la longueur limite de Blech.

2.3.1.3 Impact du choix du matériau et de la génération technologique

Les interconnexions en aluminium comportent de nombreux grains dans leur structure. Expérimentalement, il est montré que le mécanisme de diffusion prépondérant réside à *la frontière des grains* (en rouge sur la Figure 2.3). En revanche, les métallisations en cuivre ont une largeur et une épaisseur de ligne sensiblement égale aux grains les constituant **[WON 07]**. Ce type de structure est communément appelé structure bambou en référence à la structure sectionnée de la plante. Aussi, les mécanismes de diffusion de matière dans les connexions en cuivre s'en trouvent modifiés. Les chemins de diffusion les plus rapides interviennent aux interfaces *métal/barrière de diffusion* et *métal/couche de passivation* (en bleu sur la Figure 2.3) **[HU 06], [BAO 03]**.



(b) dans les lignes de cuivre.

De cette manière, la nature et les dimensions des barrières de diffusion et de la couche de passivation doivent être pris en considération. En l'occurrence, l'expression de la dérive de grain aux interfaces est directement proportionnelle au ratio entre les différents volumes. Ainsi, les interconnexions en aluminium et en cuivre peuvent être modélisées selon la Figure 2.4 où les volumes des couches de passivation V_{CP} et des barrières de diffusion V_{BD} s'expriment par les équations (2.8) et (2.9) :

$$V_{CP} = \delta_{CP} \cdot w \cdot L \tag{2.8}$$

$$V_{BD} = \delta_{BD} \cdot (2 \cdot h + w) \cdot L \tag{2.9}$$

Les expressions de la dérive de grain pour l'aluminium et pour le cuivre peuvent être réécrites respectivement par (2.10) et (2.11):

$$V_{EM_{Al}} = \frac{e \cdot \rho \cdot j}{k \cdot T} \cdot Z^*_{GRAIN} \cdot D_{0_GRAIN} \cdot e^{\frac{-E_{A_GRAIN}}{k \cdot T}} \cdot \frac{\delta_{CP}}{d} \cdot \left(1 - \frac{d}{w}\right)$$
(2.10)

$$V_{EM_{Cu}} = \frac{e \cdot \rho \cdot j}{k \cdot T} \cdot D_{0_INT} \cdot e^{\frac{-E_{A_INT_}}{k \cdot T}} \cdot \left(\left[\delta_{CP} \cdot Z_{CP}^* \cdot \left(\frac{1}{h}\right) \right] + \left[\delta_{BD} \cdot Z_{BD}^* \cdot \left(\frac{2}{w} + \frac{1}{h}\right) \right] \right) \quad (2.11)$$

où δ_{CP} et δ_{BD} sont les épaisseurs de la couche de passivation et des barrières de diffusion, d est le diamètre du grain de métal, w et h sont respectivement la largeur et la hauteur de la ligne de métal, Z^*_{GRAIN} , Z^*_{CP} et Z^*_{BD} les numéro atomiques effectifs du grain de métal, de la couche de passivation et des barrières de diffusion, $D_{0_{GRAIN}}$ et $D_{0_{INT}}$ les constantes de diffusion au niveau des gains et aux interfaces, $E_{A_{GRAIN}}$ et $E_{A_{INT}}$ les énergies d'activation de diffusion respectives à la frontière des grains et aux interfaces.



Figure 2.4: Modélisation (a) d'une interconnexion en aluminium (b) en cuivre avec une barrière de diffusion TaN/Ta.

2.3.1.4 Modèle de dégradation

Le modèle de durée de vie utilisé pour traduire le phénomène d'électromigration est le modèle semi-empirique de Black **[BLA 67]**. Il traduit le temps médian MTTF où 50% de la population des échantillons testés ont atteint la défaillance exprimé en fonction de la densité de courant J et de la température T selon l'expression (2.12) :

$$MTTF = A \cdot J^{-n} e^{\frac{E_A}{k \cdot T}}$$
(2.12)

où classiquement n \cong 2 pour les lignes AlCu, n \cong 1 pour Cu et E_A l'énergie d'activation du mécanisme, avec E_a= 0,7 eV (AlCu) et 0,9 eV (Cu). A est un paramètre considéré constant dépendant de la microstructure qui avec les équations (2.3) et (2.5) dépend de la résistivité, de la charge effective (Z*) et du coefficient de diffusion D₀.

De plus, la densité de courant peut être définie à l'aide de la tension de polarisation de la ligne ainsi qu'avec la fréquence d'émission des signaux, avec l'expression (2.13):

$$J = f \cdot \varepsilon_0 \cdot \varepsilon_{0X} \cdot \frac{V}{d} \tag{2.13}$$

L'équation de Black est une relation simplifiée des équations (1.10) et (1.11). Elle montre que les hautes températures ainsi qu'une forte densité de courant abaissent la durée de vie des lignes d'interconnexions. Cependant le modèle de Black n'est plus adapté aux technologies cuivre actuelles. Une modélisation du temps de défaillance fondée sur les mécanismes physiques de la défaillance dans les nouveaux matériaux en tenant compte des barrières de diffusion est devenue nécessaire.

Les valeurs de l'exposant n de la littérature sont présentées dans le Tableau 2.1. La valeur est sensible à la nature du conducteur, il est égal à 2 pour des interconnexions en aluminium et proche de 1 pour les interconnexions en cuivre. Cette différence de valeur peut être expliquée par la microstructure de la ligne. Si le mécanisme de nucléation de lacune domine en volume alors le coefficient n est proche de 2. En revanche si c'est le mécanisme d'extension de lacunes aux interfaces qui domine, alors n est proche de 1 (cas de la structure bambou) **[LLO07]**. Au second ordre, la valeur de n varie légèrement avec la température et la longueur de la ligne.

Valeur de n	Structure
2,02	Ligne en aluminium à T = 140 °C, W=700 nm, h = 530 nm, J = 2 à 5,4 MA/cm ² [BIE 04]
2,13	Ligne en aluminium à T = 200 °C, W=700 nm, h = 530 nm, J = 2 à 5,4 MA/cm ² [BIE 04]
1,40	Ligne-via en cuivre dual damascène 90 nm (up stream) à T = 300 °C, W = 140 nm, L = $50 \mu m$, J = 1,6 MA/cm ² , Ta/TaN [LIN 07]
1,36	Ligne-via en cuivre dual damascène 90 nm (up stream) à T = 300 °C, W = 140 nm, $L = 400 \ \mu m$, $J = 1.6 \ MA/cm^2$, Ta/TaN [LIN 07]

Tableau 2.1: Valeurs de n pour différentes structures en aluminium et en cuivre

Les valeurs des énergies d'activation sont sensibles à la nature des couches de passivation et des barrières de diffusion **[HU 03]**, **[HU 06]**. Une couche de passivation en CoWP permet par exemple de limiter le mécanisme d'électromigration à l'interface Cu/diélectrique. Les valeurs de la littérature pour les différents matériaux sont présentées dans le Tableau 2.2.

Valeurs d'E _A (eV)	Structure
0,85	Ligne en cuivre W = 130 nm + couche de recouvrement en SiN_X [HU 06]
1,10	Ligne en cuivre W = 130 nm + couche de recouvrement en $SiC_XH_YN_Z$ + barrières de diffusion en Ta/TaN [HU 06]
1,40	Ligne en cuivre W = 180 nm + couche de recouvrement et barrière de diffusion en en Ta/TaN [HU05]
1,90 - 2,40	Ligne en cuivre $W = 80$ nm (structure bambou) + couche de recouvrement en CoWP + barrières de diffusion en Ta/TaN [HU 06]
1,00	Ligne en cuivre W = 2 μ m + couche de recouvrement en CoWP + barrières de diffusion en Ta/TaN [HU 03]

 Tableau 2.2: Valeurs d'énergie d'activation traduisant l'influence des matériaux utilisés pour les barrières de diffusion

2.3.2 La migration atomique induite par une contrainte hydrostatique

Ce mécanisme est plus connu sous le nom de *Stress Induced Voiding (SIV)* ou encore *Stress Migration (SM).* Tout comme l'électromigration, c'est un mécanisme de défaillance qui se produit au niveau des métallisations du circuit intégré et qui se traduit par une augmentation de la résistance de la piste lorsque celle-ci est soumise à une contrainte hydrostatique σ_{H} . Dans le pire des cas, le mécanisme peut conduire jusqu'à un circuit ouvert. Toutefois, si le site et le mode de défaillance sont similaires à l'électromigration, le mécanisme physique ainsi que les facteurs l'aggravant sont différents. Le mécanisme de dégradation SIV est composé de 2 phases:

- La phase de nucléation de lacune,
- La phase de migration de grain et de lacune.

2.3.2.1 Physique de la défaillance

La première phase se traduit par l'apparition de lacunes au sein des joints de grains (Figure 2.5). Ce processus est conduit par l'inhomogénéité de la microstructure. A haute température et en absence de courant, les grains se dilatent et provoquent intrinsèquement des cavités dans la structure (Figure 2.5). Une étude représentant des structures cuivre double damascène d'une largeur de 0,18 µm vieillies en température, a montré que la dégradation maximale intervient pour une plage comprise entre 150°C et 200°C avec un pic critique à 190°C **[OGA 02]**.



Figure 2.5: (a)-(b) Vues FIB/MEB d'ébréchures et encoches sur les lignes métalliques (notches, mouse bites) (c) Formation de cavité (void) dans un via (source STMicroelectronics **[VIN 07]**)

La seconde phase traduit la migration et l'accumulation de lacunes dans les zones de moindres contraintes. La couche de passivation ainsi que les barrières de diffusion exercent intrinsèquement une contrainte mécanique sur le métal notamment au niveau des zones de contact via-ligne (estimé à 500 MPa pour une interconnexion de génération 0,35 µm en cuivre **[OGA 02]**). Lorsque les grains se dilatent sous l'effet de la température, il se produit alors une migration de ces mêmes grains provenant des zones les plus mécaniquement contraintes, en l'occurrence les vias vers les zones de moindre pression. En conséquence, ce retrait de matière va engendrer de manière locale une diminution de la pression et puis très rapidement l'apparition et la convergence de lacunes à l'intérieur et en dessous des vias de connexion comme le montre la Figure 2.5 c.

2.3.2.2 Modèle de vieillissement associé

2.3.2.2.1 Modèle de McPherson and Dunn

La première modélisation du mécanisme de dégradation SIV est issue des travaux *de McPherson et Dunn* en 1987 **[MCP 87]**. Ce modèle exprime la vitesse de migration des lacunes V_{R} en fonction de la température et est décrit par l'expression (2.14):

$$V_R = C \cdot (T_0 - T)^N \cdot e^{\frac{-Q_{EFF}}{k \cdot T}}$$
(2.14)

où C est une constante, T_0 la température d'absence de contrainte du matériau (°K), T la température d'application (K), N le coefficient de fluage et Q_{EFF} l'énergie d'activation de la création de lacune (eV).

Le Tableau 2.3 rassemble les paramètres de l'équation (2.14), pour une structure en aluminium et pour une métallisation cuivre de type bambou.

Matériau	С	T ₀ (K)	Ν	Q (eV)
Al	Х	507	2,33	0,58
Cu	1,5	545	3,2	0,6
77.1.1	0.2 D	11 1 1		1 001

 Tableau 2.3: Paramètres pour l'aluminium et le cuivre [ZHA 06] [YU 08]

De la même façon que pour l'électromigration, le nombre de lacune devient plus important pour des structures larges comportant de nombreux grains (aluminium) et semble en revanche plus réduit pour des structures en cuivre de type bambou.

Le modèle de *McPherson et Dunn* traduit bien l'impact du choix du matériau sur la migration des lacunes. Toutefois, le modèle présente le désavantage de ne comporter aucun paramètre technologique BEOL. Ceci laisse supposer que la migration des lacunes ne dépend pas des dimensions de la ligne ni du nœud technologique, ce qui n'est pas satisfaisant du point de vue de la physique impliquée dans le mode de défaillance.

2.3.2.2.2 Modèle de Wu

Le modèle de *Wu et al* complète le modèle de *McPherson and Dunn* dans le sens où il intègre comme paramètre le gradient de contrainte hydrostatique, qui lui, dépend de la technologie **[WU 08]**. La migration des lacunes est donnée par l'expression (2.15) :

$$V_R = C' \cdot \frac{1}{T} \cdot e^{\frac{-Q_{EFF}}{k \cdot T}} \cdot max\left(\frac{\partial \sigma_H}{\partial z}\right)$$
(2.15)

Le gradient de contrainte est calculé en extrayant la contrainte hydrostatique le long du chemin le plus critique. La contrainte hydrostatique est calculée comme étant la valeur moyenne des contraintes mécaniques s'exerçant sur la métallisation dans chaque direction, selon la relation (2.16) :

$$\sigma_H = \frac{\sigma_X + \sigma_Y + \sigma_Z}{3} \tag{2.16}$$

où σ_x , σ_y et σ_z représentent les contraintes mécaniques selon les axes x, y et z.

A ce point de l'étude, il est nécessaire d'estimer les contraintes mécaniques impliquées dans les microstructures BEOL. Ces contraintes sont fortement dépendantes des procédés de fabrication des fabricants. Toutefois elles peuvent être estimées à l'aide de la théorie des éléments finis. Dans ce cas, il est possible pour une température fixée et une technologie donnée utilisant des métallisations Al, AlCu ou Cu d'extraire le gradient de contrainte et donc la vitesse de migration des lacunes. Cette approche sera développée spécifiquement dans le chapitre 3 à l'aide de simulations de structures Al et Cu dans la section 3.3.1.2.

2.3.2.2.3 Modèle en cycle thermique

La modélisation de Wu proposée dans la section précédente, présente l'avantage de tenir compte de la nature des oxydes d'isolation, des barrières de diffusion ainsi que de la couche de recouvrement. Toutefois, il est fondé sur l'idée que la diffusion de grain est maximale lorsque la température de jonction est maximale. Ceci est vrai, si les barrières de diffusion ont un comportement élastique lorsqu'une contrainte est appliquée.

En réalité, lorsque la brique BEOL est soumise à un abaissement de la température $(\Delta T < 0)$, la couche de recouvrement et les barrières de diffusion se compriment à des vitesses différentes, du fait de la différence de leur coefficient de dilatation thermique. De ce fait, il existe donc une contrainte hydrostatique négative qui s'exerce sur les parois de la métallisation.

La fluctuation entre l'augmentation de la température ($\Delta T > 0$) et son abaissement ($\Delta T < 0$) peut engendrer la fatigue des barrières, qui peuvent au cours des cycles adopter un comportement plastique (Figure 2.6). La déformation irréversible qui en découle, conduit les grains de métaux à migrer vers ces zones délaminées, favorisant ainsi le mécanisme SIV. Toutefois, il n'existe à l'heure actuelle aucune modélisation du mécanisme SIV en cyclage thermique. Cette piste reste donc à explorer.



Figure 2.6: Courbe de contrainte en fonction de la déformation lors d'un essai de traction [VIT 06]

2.3.3 Le claquage d'oxyde intermétallique

Le claquage de l'oxyde inter métallique est le troisième mécanisme de défaillance majeur de la zone BEOL. Ce mécanisme ne concerne non plus la métallisation, mais l'oxyde à faible permittivité séparant les pistes métalliques entre elles (IMD) ou entre deux niveaux de métallisations (ILD) consécutif. Lors de la propagation des différents signaux, des charges électriques résiduelles peuvent s'accumuler aux interfaces des métallisations. Sous l'action du champ électrique, les porteurs peuvent migrer vers une ligne de potentiel plus faible et traverser l'oxyde intermétallique. Ces charges électriques peuvent former au cours du temps des fuites capacitives entre les lignes (cross talk) conduisant à un retard effectif ($\tau = RC$) à la propagation des signaux. Des courants de fuite peuvent avoir lieu entre les lignes, augmentant avec l'intégration (Figure 2.7) par la diminution de l'espacement (d) et de la largeur de la ligne (W), conduisant éventuellement au claquage de l'oxyde de manière irréversible quand un chemin de percolation survient entre la cathode et l'anode faisant perdre à l'oxyde inter métallique ses propriétés isolantes.



Figure 2.7: Mécanisme de claque d'oxyde intermétallique

2.3.3.1 Physique de la défaillance

Dans le cas de la silice et des oxydes à basse permittivité (Low-k) comme le SiOCH poreux, les porteurs de charges mis en jeu dans le mécanisme de claquage sont les électrons. Le transfert de ces électrons de la cathode vers l'anode peut s'effectuer selon deux mécanismes d'injections :

L'injection Schottky ou thermoionique est rendue possible lorsque l'énergie du porteur q Φ_e est supérieure à la hauteur de la barrière cathode-oxyde q Φ_B . Dans cette configuration, le porteur de forte énergie (porteur chaud) pénètre l'oxyde en surmontant la barrière de potentiel. En traversant l'oxyde, les porteurs peuvent créer également des pièges (neutres) qui se chargent négativement dans le volume du diélectrique (Figure 2.8).



Figure 2.8: Injection thermoionique ou Schottky directe

La densité de courant Schottky J_{CS} (A/cm²) est modélisable par la relation (2.17):

$$J_{CS} = A^* \cdot T^2 \cdot e^{\frac{-q \cdot \left(\Phi_{\rm B} - \sqrt{\frac{q \cdot E}{\pi \cdot \epsilon_0 \cdot \epsilon_{0\rm X}}}\right)}{k \cdot T}}$$
(2.17)

où A* est la constante de Richardson (120 A.cm⁻².K⁻²), Φ_B le potentiel de surface métal diélectrique (V), E le champ électrique (MV/cm), ϵ_0 la constante diélectrique du vide (8,85.10⁻¹⁴ F/cm), ϵ_{0x} la constante diélectrique de l'oxyde, k la constante de Boltzmann (8,62.10⁻⁵ eV.K⁻¹) et T la température (K), q la charge de l'électron (1,6.10⁻¹⁹ C).

L'injection Poole-Frenkel se produit lorsque l'énergie du porteur est suffisante pour qu'il puisse surmonter la barrière de potentiel entre deux pièges ΔE_{PF} et ainsi se déplacer à travers l'oxyde. Dans ce cas, le porteur passe d'un piège à un autre par conduction thermoionique locale (Figure 2.9). La conduction Poole-Frenkel vient s'ajouter à la conduction Schottky pour des champs électriques supérieurs à 1,6 MV/cm [CHE 04].



Figure 2.9: Mode d'injection par effet Poole-Frenkel comme un effet thermoionique local

La densité de courant induite par l'effet Poole Frenkel J_{PF} est donnée par l'expression (2.18) où les différents paramètres ont leur signification habituelle :

$$J_{PF} \propto E \cdot e^{\frac{-q \cdot \left(\Phi_{\rm B} - \sqrt{\frac{q \cdot E}{\pi \cdot \epsilon_0 \cdot \epsilon_{0\rm X}}}\right)}{k \cdot T}}$$
(2.18)

2.3.3.2 Modèle de dégradation

Les injections d'électrons par voie thermoionique et par conduction Poole-Frenkel sont directement liées à la valeur du champ électrique existant entre les lignes métalliques. Expérimentalement, le temps médian à la défaillance montre également une dépendance en champ électrique. De ce fait, plusieurs modèles de défaillance en E, E^{0.5} et 1/E ont été développés où E ici est le champ électrique. Certains résultats montrent que le modèle en E est le plus fidèle pour décrire le mécanisme de claquage de l'oxyde **[KIM 07]**.

De plus, le temps à la défaillance est dépendant de la surface S de conduction métaldiélectrique qui suit une statistique de Poisson **[CHE 07]**. La hauteur de ligne t est fixée par le nœud technologique, par conséquent pour une génération technologique donnée la surface S n'est définie que par la longueur de la ligne de conduction L **[CIO 06]**. Enfin, le mécanisme de claquage est accéléré à haute température et suit une loi Arrhenius.

Pour une application spécifiée, le temps de défaillance au claquage $TDDB_{APPLI}$ est exprimé pour les oxydes en silice et ceux à basse permittivité par l'équation (2.19):

$$TDDB_{APPLI} = TDDB_{TEST} \cdot e^{-\gamma \cdot (E_{APPLI} - E_{TEST})} \cdot e^{\frac{E_A}{k} \cdot \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right)} \cdot \left(\frac{S_{APPLI}}{S_{TEST}}\right)^{\beta}$$
(2.19)

où TDDB_{TEST}, E_{TEST} , T_{TEST} et S_{TEST} sont respectivement le temps à défaillance, le champ électrique, la température et la surface des échantillons dans les conditions de test, γ le paramètre d'accélération en champ (cm/MV), E_A l'énergie d'activation en température du mécanisme (eV) et enfin β la pente de la distribution de Weibull.

2.3.3.3 Oxydes à basse permittivité

Le mécanisme de claquage de l'oxyde d'isolation n'est généralement pas un mécanisme de défaillance critique pour les technologies antérieures au nœud 130 nm. Ceci s'explique notamment par l'espacement interligne important et surtout par un champ de claquage intrinsèque relativement élevé de la silice (10-15 MV/cm) principalement utilisée. Plus récemment, l'introduction des oxydes à faible permittivité (Low-k) dans les générations sub 130 nm a profondément modifié l'impact de ce mécanisme sur la fiabilité du composant et cela pour plusieurs raisons :

- Les oxydes à basse permittivité possèdent une hauteur de barrière ainsi qu'un champ de claquage intrinsèque 2 à 3 fois plus faible que celui de la silice (1,5 6 MV/cm pour le SiOCH) [KIM 07] [LLO 04]. Ceci signifie que le claquage franc de l'oxyde intervient pour des tensions 2 à 3 fois plus faibles.
- Les oxydes à basse permittivité ont des propriétés mécaniques moins bonnes que leur homologue en SiO₂. Le fort coefficient d'expansion thermique des oxydes par rapport aux barrières de diffusion ainsi que leur faible Module de Young, implique la présence de contraintes résiduelles importantes aux interfaces, à faible température.
- Ces oxydes sont poreux et comportent de nombreux sites hydrophiles (SIOH, COOH) [VIA 02]. Ces sites peuvent capturer des molécules d'eau et/ou des espèces hydrogénées lors des différentes étapes de gravure et de polissage mécano-chimique (CMP) provoquant ainsi des délaminations et des corrosions locales [LLO 04] conduisant ainsi à des courants de fuite au sein de la structure.

2.4 Les mécanismes de défaillance FEOL

2.4.1 Le claquage de l'oxyde de grille

Le claquage d'oxyde de grille ou TDDB (Time Dependent Dielectric Breakdown) est une défaillance de l'oxyde de grille résultant de l'apparition de défauts générés lors d'une contrainte électrique qui provoque une injection non contrôlée de porteurs dans l'oxyde conduisant à la perte définitive des propriétés isolantes du diélectrique. Les mécanismes d'injections et la fiabilité qui en découle dépendent fortement de la nature de l'empilement (stack de grille), selon qu'il s'agit de silice (SiO₂), d'oxyde nitruré (SION) ou d'oxyde à haute permittivité (High K).

2.4.1.1 Oxydes SiO2 et SiON

2.4.1.1.1 Mécanismes d'injection dans les oxydes SiO2 et SiON

L'architecture CMOS impose aux transistors NMOS et PMOS de fonctionner soit en inversion forte, soit en régime d'inversion faible et éventuellement en faible déplétion. Ceci implique que dans les oxydes de grille épais à intermédiaire ($T_{ox} > 3nm$) l'injection de porteurs à travers l'oxyde, se produit principalement du canal vers la grille avec la tension appliquée sur la grille V_G .

En régime d'inversion, la polarisation abaisse la barrière de potentiel grille-substrat, ce qui permet aux porteurs d'énergie suffisante, de traverser l'oxyde du canal vers la grille. Dans le cas d'un transistor NMOS associé à un oxyde en silice, les porteurs injectés en inversion sont des électrons puisque la hauteur de la barrière de conduction ΔE_C est plus basse que celle de valence ΔE_V (3.1 eV contre 4.5 eV). Dans le cas du transistor PMOS en inversion, ce sont les trous qui sont injectés. Ce type de conduction est qualifié d'injections Fowler-Nordheim et s'exprime par la relation (2.20) :

$$J_{FN} = a \cdot F_{OX}^2 \cdot e^{\frac{-b}{F_{OX}}}$$
(2.20)

où J_{FN} est la densité de courant injecté, F_{OX} représente le champ électrique vertical, et a et b sont des paramètres dépendants de la hauteur de barrière effective φ_0 à l'interface semi-conducteurisolant et de la masse effective des électrons (m_{ox}^*) dans la bande de conduction de la silice.

Toutefois, les transistors des générations CMOS fortement submicroniques $L_G \leq 0,13 \mu m$ HP et LP disposent d'oxydes de grille dont l'épaisseur physique est inférieure à 3 nm. Prenons l'exemple de l'inverseur CMOS (Fig. 2.10 à 2.13) où le transistor PMOS de charge (Pull Up) est polarisé avec la tension d'alimentation (V_{DD}) et le transistor NMOS de signal (Pull Down) qui est lui polarisé à la tension V_{SS} . Pour ces épaisseurs, les oxydes de grille sont sujets aux injections de porteurs même à faible champ (régime sous seuil $V_{GS} < V_{TH}$, $V_{OUT} = V_{DS} = V_{DD}$), en mode d'injection Tunnel (Figure 2.10) (Figure 2.12). Pour les cellules CMOS, les fuites par effet tunnel viennent s'ajouter aux porteurs injectés par injections Fowler-Nordheim en régime inversion (Figure 2.11) (Figure 2.13) quand $V_{GS} = V_{In}$ augmente à V_{DD} (V_{Out} va alors vers l'état bas).


2.4.1.1.2 Physique de la défaillance

Le modèle traditionnellement utilisé pour modéliser le claquage de l'oxyde de grille (SiO₂) est le modèle de percolation (Figure 2.14). Ce modèle reste également valide pour les oxydes nitrurés (SiON). Lors d'une excitation électrique, les porteurs injectés vont casser les liaisons à l'interface substrat-isolant (états d'interface) occasionnant tout d'abord la dépassivation des liaisons Si-H (H release) et vont éventuellement générer des piégées aléatoirement dans le volume du diélectrique de grille. L'accumulation de ces pièges à travers l'oxyde permet la création d'un chemin de conduction par un effet de percolation, permettant ainsi l'apparition d'un courant de fuite croissant avec le temps et le nombre de défauts à travers la structure. Le claquage intervient lorsqu'une densité critique de pièges est atteinte conduisant à la perte totale des propriétés isolantes du diélectrique **[WU 05]**.



Figure 2.14: Modèle schématique de conduction par effet de percolation [WU 05]

L'accumulation de pièges décrit par le modèle de percolation, traduit à la fois un mécanisme d'usure progressive et de défaillance brutale du diélectrique. Toutefois, les oxydes minces $T_{ox} < 5$ nm subissent un nombre moins important de pièges au claquage à travers l'oxyde. Ceci est dû en premier lieu au fait que la somme de la distance tunnel et du libre parcours moyen des porteurs (1,5 à 2 nm) deviennent de l'ordre de l'épaisseur d'oxyde. Ceci se traduit par une diminution de la pente de la distribution de Weibull, β , avec la réduction de T_{ox} dans les générations fortement submicroniques. Si l'on suppose un diamètre de défaut a_0 constant, cette tendance est exprimée par la dépendance de la pente β par la relation (2.21) **[WU 05]** :

$$\beta \approx \frac{T_{OX}}{a_0} \tag{2.21}$$

Les faibles valeurs de β (aux alentours de 1 pour un oxyde d'épaisseur physique 1,5 nm) montrent que le TDDB devient alors un mécanisme aléatoire actif durant la période de vie utile qui affecte principalement le taux d'apparition de la défaillance λ (= 1/T_{DDB}) et non plus systématiquement la durée de vie (Figure 2.15).



Figure 2.15: (a) Dépendance linéaire de la pente de Weibull avec l'épaisseur (physique) d'oxyde de grille, (b) Distribution de Weibull du temps de défaillance TDDB pour différentes épaisseurs d'oxyde **[WU 05]**.

2.4.1.1.3 Modèles de défaillance pour les oxydes SiO2 et SiON

Expérimentalement, il est montré que le temps au claquage du diélectrique de grille dépend essentiellement de 4 paramètres :

- L'épaisseur physique d'oxyde T_{ox} [WU 05],
- Les surfaces d'oxydes A_{OX} [PIC 06] [PRE 98],
- Les tensions appliquées sur la grille Vg [WU 05],
- La température T [WU 05].

Les deux modèles classiquement utilisés pour caractériser la dépendance en champ électrique du claquage de l'oxyde épais ($T_{ox} > 7$ nm) sont les modèles dits « thermochimiques » (modèle en E) et le mécanisme dit « Anode-Hole-Injection » (AHI) (modèle en 1/E).

Le modèle en E considère que la génération de défauts est provoquée principalement par l'amplitude du champ électrique E et l'effet thermique impliqué dans la rupture de liaisons atomique et moléculaire comme défauts précurseurs du claquage (lacunes d'oxygène, liaisons Si-O-Si, Si-Si). Le courant traversant l'oxyde joue un rôle secondaire [MCP 98]. Le temps de claquage médian TDDB_{APPLI} est donné dans les conditions d'applications par la relation (2.22) :

$$TDDB_{APPLI} = TDDB_{TEST} \cdot e^{-\gamma \cdot (E_{APPLI} - E_{TEST})} \cdot e^{\frac{E_A}{k} \cdot \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right)} \cdot \left(\frac{S_{APPLI}}{S_{TEST}}\right)^{\frac{1}{\beta}}$$
(2.22)

où TDDB_{TEST}, E_{TEST}, T_{TEST} et S_{TEST} sont respectivement le temps à défaillance, les champs électriques, la température et la surface de référence, γ le paramètre d'accélération en champ (cm/MV), E_A l'énergie d'activation en température du mécanisme (eV), S la surface de référence (cm²) et enfin β la pente de la distribution de Weibull.

Le modèle en 1/E décrit le mécanisme AHI dans lequel les électrons, accélérés par le champ vertical, sont injectés depuis l'oxyde et perdent leur énergie à l'interface substrat-oxyde par ionisation par impact dans l'anode où ils créent des paires électrons trous. Ces trous énergétiques générés dans l'anode sont suffisamment efficaces pour générer des pièges dans l'oxyde par rétro-injections jusqu'à la cathode, favorisant ainsi l'apparition d'un chemin de percolation. Le temps de défaillance du modèle est donné dans les conditions d'applications par l'expression (2.23) :

$$TDDB_{APPLI} = TDDB_{TEST} \cdot e^{G \cdot \left(\frac{1}{E_{APPLI}} - \frac{1}{E_{TEST}}\right)} \cdot e^{\frac{E_A}{k} \cdot \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right)} \cdot \left(\frac{S_{APPLI}}{S_{TEST}}\right)^{\frac{1}{\beta}}$$
(2.23)

où G est le facteur d'accélération en champ ($MV.cm^{-1}$) et E_a l'énergie d'activation en température du mécanisme (eV).

Pour des oxydes ultrafins $T_{ox} < 3$ nm, les modèles précédents ne sont plus valables puisque le champ électrique important est causé non pas par une polarisation importante mais par une épaisseur d'oxyde faible **[WU 05]**. Wu a montré que le mécanisme de claquage présentait pour ces épaisseurs d'oxyde, une forte dépendance en tension mais surtout un comportement qui ne suit plus une loi d'Arrhenius en température. Ceci signifie que l'énergie d'activation doit être réexaminée pour les oxydes d'épaisseur inférieure à 3 nm notamment en incluant la dépendance en champ du facteur d'accélération **[VIN 96]**. Dans les conditions d'application, le temps de claquage médian a été proposé comme suivant la relation (2.24) **[WU 05]**:

$$TDDB_{APPLI} = TDDB_{TEST} \cdot \left(\frac{V_{GS_APPLI}}{V_{GS_TEST}}\right)^{-N} \cdot e^{X\left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right) + Y\left(\frac{1}{T_{APPLI}^2} - \frac{1}{T_{TEST}^2}\right)} \cdot \left(\frac{S_{APPLI}}{S_{TEST}}\right)^{\frac{1}{\beta}}$$
(2.24)

où V_{GS_APPLI} et V_{GS_TEST} sont respectivement les tensions grille-source dans les conditions d'application et dans les conditions de test (V). X et Y sont des constantes d'ajustements (estimées respectivement à 8,81.10³ K et à – 7,75.10⁵ K²), N est une constante dépendant de la nature de l'oxyde.

2.4.1.2 Oxydes à haute permittivité HK

Les modèles présentés dans le paragraphe précédent ne sont applicables que pour des $oxydes SiO_2$ ou SiON. A partir de la génération CMOS 45 nm HP, certains fabricants ont commencé à remplacer l'oxyde de grille par un empilement constitué d'un oxyde à haute permittivité (Hi-K) et d'une couche d'interface (IL) d'oxyde. Classiquement pour le nœud 45 nm, l'oxyde de grille à haute permittivité est un dérivé de hafnium (~ 3 à 6 nm) alors que la couche d'interface est réalisée à partir de silice amorphe dont l'épaisseur physique est proche du nanomètre **[CHO 08]**, **[POR 06]**.

La compréhension des mécanismes de défaillance impliqués dans les empilements formés de l'oxyde à haute permittivité et de l'oxyde d'interface, nécessite de prendre en compte plusieurs paramètres. Le premier paramètre concerne la valeur des champs de claquages intrinsèques des oxydes de l'empilement. Ces valeurs déterminent les tensions de grille critiques pour lesquelles le claquage franc de l'empilement peut se produire (HBD ou hard breakdown). Les champs électriques sont estimés à 5 MV/cm pour un oxyde HfO_2 de 5 nm et de 15 MV/cm pour en oxyde en silice de 1 nm [CHO 08].

Le deuxième paramètre concerne les épaisseurs physiques des oxydes de l'empilement. Le claquage franc n'est pas le seul indicateur de la défaillance. Pour les empilements Hi-K et plus généralement pour les technologies CMOS inférieures à 0,13 µm, il est impératif de considérer les claquages doux (SBD) et le claquage progressif (PB) comme phénomènes précurseurs au claquage franc. Pour cela, il est nécessaire de comprendre les mécanismes d'injections et de connaitre les épaisseurs physiques des différents oxydes de l'empilement.

2.4.1.2.1 Mécanismes d'injections dans les oxydes à haute permittivité

Les mécanismes d'injections sont directement reliés aux épaisseurs des différents oxydes de l'empilement. S'il est acquis que toute conduction à travers la couche d'interface de 1nm se fait par effet tunnel direct, le mécanisme d'injection à travers l'oxyde à forte permittivité varie selon son épaisseur physique.

Pour une épaisseur supérieure à 3,3 nm (CMOS 45 nm HP), la conduction est de type Poole-Frenkel (Figure 2.16), (Figure 2.17). Sous l'effet de la polarisation positive de la grille, les charges électriques peuvent traverser l'oxyde Hi-K jusqu'à l'anode, en sautant de piège en piège par conduction thermoionique locale (régime AC). L'origine de ces charges dans l'oxyde Hi-K est toujours sujette à discussion. Des études récentes montrent trois origines différentes :

- Une grande quantité de charges injectées provient de la cathode, qui en régime d'inversion peuvent traverser l'oxyde d'interface par effet tunnel jusqu'à atteindre l'oxyde Hi-K [YOU 06],
- Des pièges neutres sont initialement présents dans le volume de l'oxyde HK [YOU 06].
 Les électrons injectés peuvent être piégés et dépiégés (Poole-Frenkel) et participer aux courants de fuite grille-canal. Cet effet peut de même perturber de manière réversible la stabilité de la tension de seuil (NBTI),

Enfin, la surface de contact entre l'oxyde IL à l'interface et le diélectrique Hi-K peut comporter également de nombreux défauts. Ces défauts peuvent piéger des porteurs provenant du canal en régime d'inversion et influencer la tension de seuil. Cependant, le champ appliqué à travers l'empilement ne semble pas assez important pour pouvoir dépiéger ces porteurs [CHO 08].

Pour ces oxydes épais, la défaillance de l'empilement s'effectue en deux temps :

- L'oxyde d'interface claque lorsque $E_{OX IL} > E_{OX IL-BD} = 15 \text{ MV/cm}$,
- Ensuite l'oxyde HK claque à son tour lorsque $E_{OX_{HK}} > E_{OX_{HK}-BD} = 5 \text{ MV/cm}.$

Expérimentalement, la défaillance de l'empilement n'est pas observée avant que l'oxyde Hi-K claque **[CHO 08]**. Ainsi, lorsque l'épaisseur physique de l'oxyde à forte permittivité est supérieure à 3,3 nm $T_{OX HK}$ > 3,3 nm, la fiabilité de l'empilement FEOL est déterminée par l'oxyde Hi-K.

Pour les oxydes Hi-K inférieurs à 3,3 nm (CMOS 32-22 nm), la conduction de porteurs est assurée par effet tunnel (Figure 2.18) (Figure 2.19). Sous l'effet du champ électrique, les porteurs peuvent traverser par conduction tunnel les deux oxydes de l'empilement. Aussi, la faible épaisseur de l'oxyde Hi-K minimise le chemin d'interaction entre les porteurs et les atomes de l'oxyde. De ce fait, les porteurs cèdent beaucoup moins d'énergie lors de leur transfert vers l'anode. Par conséquent au niveau de la grille, les électrons peuvent générer des paires électronstrous et provoquer par la même occasion une injection inverse de trous, décrit par le modèle AHI [KAU02]. Ce mécanisme est beaucoup plus marqué pour les oxydes Hi-K minces.





Pour ces oxydes Hi-K minces, la défaillance de l'empilement s'effectue également en deux temps :

- L'oxyde Hi-K claque en premier lorsque $E_{OX_{HK}} > E_{OX_{HK}_{BD}}$ (Figure 2.20),
- Ensuite l'oxyde d'interface claque à son tour lorsque $E_{OX_{IL}} > E_{OX_{IL}BD}$ (Figure 2.20)

Expérimentalement, aucune défaillance de l'empilement n'est observée avant que l'oxyde d'interface claque **[POR 06]**. De ce fait, lorsque l'épaisseur physique de l'oxyde de grille à forte permittivité est inférieure à 3,3 nm ($T_{OX_{HK}} < 3,3$ nm), la fiabilité de l'empilement FEOL est déterminée principalement par l'oxyde de la couche IL à l'interface **[POR 06]**.



Figure 2.20: Caractéristique I_G-V_G d'un empilement oxyde $HfO_2/SiO_2 = 2.5/0,6$ nm [POR 06]

2.4.1.2.2 Modèle de défaillance pour les oxydes à haute permittivité

Les empilements Hi-K sont sensibles aux mêmes paramètres que les oxydes SiO2, c'est-à-dire :

(i) pour les paramètres physiques ; l'épaisseur et la surface de l'échantillon sous contrainte [RIB 07],

(*ii*) pour les paramètres expérimentaux [MCP03] ; la température [MCP 03]; l'amplitude et le sens de la polarisation électrique [CHO 08].

Il n'existe pas à l'heure actuelle de loi de vieillissement simple sur l'empilement (Hi-K+SiO₂) traduisant le temps de défaillance en fonction des différents paramètres physiques et expérimentaux. En revanche, le modèle thermochimique représentatif des oxydes en silice épais (2.25), a été adapté pour les oxydes à haute permittivité seuls [MCP 02], [MCP 03]. Le modèle est adapté au Hi-k avec γ , E_a et β suivant la relation :

$$TDDB_{APPLI} = TDDB_{TEST} \cdot e^{-\gamma \cdot (E_{APPLI} - E_{TEST})} \cdot e^{\frac{E_A}{k} \cdot \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right)} \cdot \left(\frac{S_{APPLI}}{S_{TEST}}\right)^{\frac{1}{\beta}}$$
(2.25)

Oxyde Hi-K seul	ε	EOT (nm)	γ (MV/cm)	E _A (ev)	β	
HfO ₂	25	1,35	13,2	1,6	1,3 – 1,6	
HfSiON	10	2.3	6.5	0,9	1.4	
Ta_2O_5	26	22	13,6	1,7	1,5	

Les valeurs des paramètres du modèle sont consignées dans le Tableau 2.4.

 Tableau 2.4: Paramètres du modèle thermochimique HK [MCP 03] [KIM 03]

Les valeurs des pentes des distributions de Weibull pour différents oxydes Hi-K ($\beta = 1,3$ à 1,6) sont légèrement différentes de celles de l'oxyde SiO₂ et proches de celle de la génération CMOS 65 nm (Figure 2.15). Suñé a proposé un modèle analytique simple [SUN 01] dérivé du modèle de percolation où il a reformulé l'eq. (2.19) comme $\beta = (T_{ox}+T_{inter})/a_o$ où T_{inter} est l'épaisseur empirique équivalente à l'interface. Cependant il faut souligner que le claquage des diélectriques Hi-K est fortement dépendant des différentes épaisseurs respectives (Hi-K + IL) et des modes d'injections (électrons/trous) ainsi que de la polarité. Par exemple dans des oxydes de 3nm d'HfO₂, (EOT = 1,35 nm), le claquage s'apparente à un claquage progressif observé par des sauts en courant par une dégradation progressive du diélectrique qui peut être précurseur du claquage franc [GAR 04]. Le claquage progressif est généralement une des caractéristiques des oxydes SiO₂

ultra-minces avec le mécanisme de fuites de type Stress Induced Leakage Current (SILC). Cependant, à ce stade de l'étude et pour ce qui nous intéresse, c'est à dire modéliser simplement le TDDB pour les oxydes Hi-K, il n'est pas nécessaire de détailler plus avant ces différentes contributions. Néanmoins, il est important de souligner que le claquage des oxydes à haute permittivité est également un mécanisme de dégradation sévère pouvant limiter la durée de vie des dispositifs.

2.4.2 Les porteurs chauds

La dénomination de porteurs "chauds" désigne les porteurs de charges ayant gagné suffisamment d'énergie (cinétique) sous l'effet du champ électrique pour enclencher le phénomène d'ionisation par impact et générer des défauts à l'interface et dans le diélectrique de grille. Ce mécanisme d'injection peut être utilisé dans le cadre de la programmation ou l'effacement des mémoires non volatiles de type flash ou EEPROM **[MAS 07]**. Dans le cas des circuits CMOS, le phénomène porteurs chauds est néfaste au bon fonctionnement des cellules et se traduit par une dérive nette des paramètres du transistor représentatifs des performances en vitesse (G_m , ID_{sat}, $\mu_{n,p}$) et en stabilité des états On/Off (VT_{n,p}, Ion/Ioff) lors de la commutation. C'est pourquoi la fiabilité ou dégradation "porteurs chauds HCI" (*Hot carrier injection*) est considérée comme un mécanisme d'usure représentatif de la qualité d'un nœud technologique et précurseur de l'apparition des modes de défaillance plus sévère. La dégradation par porteurs chauds traduit ainsi, par abus de langage, la résistance au vieillissement progressif des transistors et des cellules numériques en restant associée aux modes de défaillance. Elle est généralement associée à la fiabilité globale d'une technologie, devenant un standard d'évaluation de la qualité des procédés de fabrication des circuits avancés.

La dégradation porteurs chauds HCI se traduit par un mécanisme d'usure lié à la dégradation progressive de l'oxyde de grille et de son interface avec le substrat au cours de la commutation AC des transistors de type MOSFET. Au niveau physique, ce mécanisme conduit à la formation de défauts à l'interface (N_{TT} +/- Donneur, Accepteur) oxyde-substrat ainsi qu'à la génération et/ou remplissage de pièges dans le volume de l'oxyde (N_{OT} +/-). Les états d'interface et les pièges d'oxyde sont directement à l'origine de l'augmentation de la tension de seuil $V_{TH}(t)$ du transistor au cours du temps selon l'expression :

$$V_{TH}(t) = V_{TH}(t_0) + \frac{q \cdot N_{IT}(t)}{C_{OX}} + \frac{q \cdot N_{OT}(t)}{C_{OX}}$$
(2.26)

où $V_{TH}(t_0)$ est la tension de seuil initiale du transistor, C_{OX} la capacité de l'oxyde de grille par unité de surface (F/cm²) et où ici les défauts sont chargés négativement (cas du transistor NMOS).

Lors des phases d'accélération des mécanismes de dégradation, par exemple dans un transistor NMOS standard, les paramètres tels que le courant de saturation en mode statique (DC) et la fréquence de fonctionnement (AC) se trouvent diminués avec le temps de contrainte. L'importance de la dégradation porteurs chauds est fondamentalement dépendante de l'optimisation du nœud technologique (L_G , T_{ox} , V_{DD}), des mécanismes d'injections, des conditions de stress et du type de transistor, qu'il est nécessaire de détailler dans les sections suivantes.

2.4.2.1 Physique de la défaillance pour les oxydes SIO2, SION et les oxydes à haute permittivité

Dans le cas d'un transistor NMOS, les porteurs chauds mis en jeu dans la dégradation sont des électrons du canal d'inversion alors que pour un transistor PMOS, ce sont les trous. L'origine de la dégradation par porteurs chauds est relative à l'enclenchement du phénomène d'ionisation par impact que nous décrirons brièvement. Les sections suivantes présenteront le cas du transistor NMOS et PMOS où nous allons distinguer principalement deux régimes porteurs chauds suivant :

- les épaisseurs d'oxyde de grille intermédiaire à épais $T_{ox} > 3$ nm
- les nœuds technologiques fortement submicroniques 1,5 nm \leq T_{ox} \leq 3 nm

2.4.2.1.1 Phénomène d'ionisation par impact (IP)

Les porteurs minoritaires sont accélérés jusqu'au point de pincement du canal grâce à la différence de potentiel appliquée entre la source et le drain. Sous l'action du champ électrique latéral existant entre le point de pincement et le drain, les électrons peuvent alors être injectés dans la zone désertée drain-substrat et, si leur énergie est suffisante, peuvent ioniser les atomes du substrat par collision inélastique en générant des paires électron-trou.

Ce phénomène de transfert d'énergie par collision inélastique, appelé ionisation par impact, permet de générer des paires électrons-trous dans la zone de charge d'espace au voisinage du drain (Figure 2.21). Cette énergie va donc permettre à un électron de la bande de valence d'être excité et passer dans la bande de conduction laissant un trou dans la bande de valence. L'énergie totale et le moment cinétique du porteur sont conservés lors du processus et les électrons fortement énergétiques sont injectés suivant la distribution des champs électriques vers la grille: L'efficacité d'injection est fortement dépendante du champ électrique vertical E_{OX} avec V_{GS} , présent entre la grille et le substrat. Si l'énergie des électrons est suffisante, *i.e.* supérieure à la hauteur de barrière de conduction de l'oxyde, ($\phi_{Bo,e} \sim 3,2$ eV pour un oxyde SiO₂), alors les électrons peuvent traverser l'oxyde de grille près de la région du drain [**DIG 06**]. Le courant injecté mesuré à la grille (I_G) sera donc représentatif de l'efficacité d'injection des porteurs suivant la tension V_{GS} et l'épaisseur d'oxyde de grille.



Figure 2.21: Mécanisme d'ionisation par impact dans un transistor NMOS en mode saturé

Conjointement au phénomène d'ionisation, quand les électrons sont injectés en direction du drain et de l'oxyde de grille, les trous sont repoussés vers le substrat. Il apparait alors un courant (négatif) de trous dans le substrat I_{SUB} dans le transistor NMOS (Figure 2.21). Ce courant est un bon indicateur de l'intensité du phénomène d'ionisation et fort utile pour l'analyse de la dégradation du transistor par porteurs chauds car il est directement dépendant du pic de champ latéral au drain (2.27). D'autre part il est communément admis qu'il y a pratiquement autant de paires électron-trous générées que de trous collectés par le substrat. La modélisation des électrons chanceux (*Lucky Electron Model* ou LEM) est représentée par le courant de substrat donnée par l'expression (2.27) **[HU 85]** :

$$I_{SUB} = C_1 \cdot (V_{DS} - V_{Dsat}) \cdot I_{DS} \cdot e^{\left(\frac{-\Phi_{i,e}}{q \cdot \lambda_e \cdot E_{LAT}}\right)}$$
(2.27)

où C_1 est une constante, (V_{DS} - V_{Dsat}) la tension tenue par la zone de charge d'espace au drain, I_{DS} le courant de drain-source, $\Phi_{i,e}$ l'énergie d'ionisation (~1,3 eV) **[TAM 84]**, q la charge de l'électron, λ_e le parcours libre moyen et E_{LAT} , le champ électrique latéral.

Dans les oxydes intermédiaires à épais, le courant substrat est maximal pour $V_{GS} \cong V_{DS}/2$, ce qui constitue la condition de polarisation pire cas standard pour la génération de porteurs chauds du canal des technologies $V_{DD}=2,5$ V à 5 V avec $T_{ox}=5$ nm à 20 nm [**BRA 03b**]. La première conséquence est la génération d'une densité importante d'états d'interface à proximité du drain qui peut s'étendre vers le canal avec le temps de contrainte (Figure 2.21) suivant la polarisation en V_{GS} par rapport à V_{DS} . Cette représentation correspond à celle de "l'électron chanceux" décrite par Hu *et.al* en 1985 [**HU 85**].

De plus, contrairement aux mécanismes de défaillance du silicium, il faut souligner que le phénomène d'ionisation augmente lorsque la température de l'environnement d'utilisation diminue. En effet, les basses températures réduisent l'agitation thermique des atomes du substrat. Ceci conduit à ce que la probabilité de collision entre les porteurs libres et les atomes du réseau diminue, ce qui a pour origine la réduction de l'interaction électron phonons. De ce fait, le libre parcours moyen des porteurs (λ_e , λ_h) augmente dans la zone de charge d'espace, ce qui permet aux porteurs d'acquérir une plus forte énergie accélérés par le champ, avant une éventuelle collision avec le réseau. Lorsque la collision a lieu, les porteurs de charges cèdent davantage d'énergie au réseau en se thermalisant dans la bande de conduction et la création de paires électron-trou devient plus importante. Ceci se traduit par l'augmentation du courant I_{SUB} à basse température (équation 2.25) étant donné que l'énergie thermique est directement reliée au champ latéral $k_B T_{eff} = q.\lambda_e.E_{LAT}$. A l'inverse, le phénomène d'ionisation par impact est fortement réduit à haute température, ce qui montre que pour les technologies à oxydes de grille épais, la fiabilité porteurs chauds est sensiblement meilleure pour des températures de mission importantes qu'en ambiant ou à basse température. Ce résultat est un comportement opposé à la dégradation de type NBTI qui augmente plus fortement avec la température notamment dans les transistors PMOS que nous traiterons spécifiquement dans la section 2.4.3.

2.4.2.1.2 Mécanismes d'injections de porteurs chauds du canal (CHC)

Le mécanisme de dégradation par porteurs chauds a pour origine l'injection de porteurs vers l'oxyde de grille directement à partir de la zone grille-drain où est localisé le phénomène d'ionisation par impact et la génération maximum de paires électron-trou. Ce mécanisme est appelé porteurs chauds du canal ou CHC (*Channel Hot Carrier*) car il correspond au domaine d'accélération en polarisation modérée avec le champ latéral telle que $V_{DS} = 1,1$ à 1,35 V_{DD} en régime saturé, c'est à dire pour $V_{GS} > V_{TH}$ et $V_{DS} > V_{DSAT}$. Le courant injecté de porteurs chauds occasionne des défauts ponctuels à l'interface et dans l'oxyde de grille conduisant à la dérive des caractéristiques électriques comme le courant linéaire, saturé et les gains du transistor. Les porteurs peuvent entrer en collision et casser les liaisons $Si_3 \equiv H$ à l'interface quand leur énergie Φ_{TT} atteint 3,7 eV. Cette valeur correspond à la valeur de la hauteur de barrière de l'oxyde additionnée à l'énergie nécessaire qu'il faut pour casser la liaison covalente $Si_3 \equiv H$ **[HU 85]**. Ce mécanisme dominant pour les stress à la polarisation $V_{GS} = V_{DS}/2$ génère ainsi des liaisons pendantes à l'interface oxyde-substrat qui vont se charger négativement par les électrons du canal en inversion (NMOS) remplissant les états d'interface N_{TT} , lors du fonctionnement nominal (Figure 2.21). Ces états d'interface réduisent la mobilité (μ_n) des porteurs du canal et par conséquent diminuent les performances en courant du transistor. A ce point de l'étude, il est nécessaire de distinguer les principales influences technologiques vis-à-vis des dégradations prédominantes sous injections de porteurs chauds du canal en fonction des polarisations V_{GS} , V_{DS}

2.4.2.1.2.1 Génération technologique à oxyde de grille épais $T_{ox} > 3$ nm

Le mécanisme pire-cas de dégradation dans le NMOS à oxyde de grille épais à intermédiaire est généralement attribué à la génération des états d'interface accepteur (ΔN_{TT}) en corrélation au maximum de courant substrat $V_{GS} \sim V_{DS}/2$, *i.e.* dû à l'injection simultanée d'électrons et de trous chauds. Ces états d'interface chargés en régime d'inversion réduisent à la fois la concentration des électrons du canal et leur mobilité effective (μ_{eff}) produisant ainsi une augmentation de la résistance du canal et la réduction du courant de drain (I_{DS}) et de la transconductance (G_m) en mode linéaire [**BRA 09**]. En raison de la localisation de la zone dégradée (Figure 2.22a, région 1), la tension de seuil V_{TH} est beaucoup moins affectée que le courant $I_{DS,lin}$ ou G_m . De plus, suivant l'extension de la région 1 dégradée, V_{TH} peut également être moins dégradé que le courant entrant en régime saturé *i.e.* pour $V_{DS} > V_{DSat}$, dû à l'exclusion d'une partie de la zone des défauts avec V_{DS} .



Figure 2.22: Mécanisme de dégradation principaux par porteurs chauds dans les nœuds technologiques à oxyde de grille épais (Tox > 3 nm) (a) Transistor NMOS, (b) Transistor PMOS.

L'amplitude de V_{GS} modifie le nombre et le type de porteurs impliqués dans la dégradation. A basse tension ($V_T \leq V_{GS} \leq V_{DS}/5$), le courant peut augmenter dû au piégeage de trous avant que la génération des états d'interface accepteurs ne redevienne le mécanisme dominant sur le courant, après neutralisation d'une partie des trous piégés [**BRA 09**]. A l'inverse, l'augmentation du champ vertical augmente fortement l'efficacité d'injection des électrons dans les régions 2 et 3 (Figure 2.23a) ce qui conduit au piégeage d'électrons ΔN_{OT} - pour les hautes tensions V_{GS} dans la région de recouvrement grille-drain (région 2) jusque dans l'espaceur d'oxyde (région 3). Ces défauts sont connus pour ajouter un effet résistance-série qui contribue à la réduction en courant du transistor NMOS.

La dégradation dans le transistor PMOS est reconnue comme moins sévère que dans le NMOS étant donné le plus faible taux d'ionisation par impact à tension fixée **[BRA 09]**. Cette différence fondamentale provient de la plus faible mobilité des trous du canal $\mu_p \sim \mu_n/3$ et du seuil d'ionisation par impact plus élevé $\phi_{i,h} = 2,54$ eV **[TZO 86]** pour enclencher la génération de paires électron-trou. D'autre part comme la hauteur de barrière (à champ nul) entre le silicium et l'oxyde de grille pour les trous $\phi_{Bo,h} = 4,8$ eV est plus élevée que celle faisant face aux électrons ($\phi_{Bo,e} = 3,2$ eV), l'injection d'électrons chauds apparait de façon importante à bas V_{GS} mais pour des conditions de champ latéral (et vertical) plus fortes. Ce régime d'injections d'électrons est proche de l'avalanche et est désigné comme le mode Drain Avalanche Hot Electron (DAHE) **[KOY 86]**.

Le mode de dégradation principal dans les technologies 0,5 μ m – 0,35 μ m est le piégeage d'électrons dans l'oxyde de grille localisé au drain (Figure 2.23b). Cette localisation de la charge d'oxyde conduit à une extension de la zone p- LDD (MDD) de la région 2 vers 3, qui entraîne une réduction significative de la longueur électrique effective du transistor avec la zone des défauts, masquant ainsi l'effet des états d'interface (donneurs) générés simultanément. Cette charge ΔN_{OT} - localisée entraîne une augmentation de la transconductance G_m et du courant | I_{DSp} | conjointement à la réduction de | V_{TH} | et de L_{EFF} . Il est alors question de l'effet de raccourcissement du canal (L_{eff} *shortenning*), ce qui explique une des raisons pour laquelle le transistor PMOS pouvait subir des effets de perçage plus prononcé avec la dégradation porteurs chauds [KOYA 86].

C'est également pour ces raisons que la dégradation du transistor PMOS a traditionnellement moins requis d'attention que le NMOS étant donné que sa dégradation se

traduisait par un gain relatif en performance qui compensait partiellement la dégradation du NMOS. Par conséquent, quand nous parlerons de dégradation porteurs chauds pire cas pour les technologies CMOS à oxyde de grille épais à intermédiaires ($T_{ox} > 3$ nm), ceci se réfèrera strictement au cas des transistors NMOS.

2.4.2.1.2.2 Génération technologique à oxyde de grille ultra-minces $1,5 \le Tox \le 3$ nm

Avec l'intégration des filières CMOS, l'optimisation du diélectrique de grille a pris de plus en plus d'importance au fur et à mesure de la réduction de Tox, LG et VDD pour les différentes filières CMOS successives. L'augmentation du champ vertical conjointement à la réduction de T_{ox} a changé la nature des courants injectés dans les oxydes ultra-minces où la grande majorité des porteurs atteignent la grille par effet tunnel direct. Il en résulte que dans les oxydes de grille tels que $T_{ox} \leq 3$ nm, la population d'électrons chauds se retrouve noyée dans la composante tunnel ce qui rend difficile la distinction du courant thermoionique de porteurs (I_G) responsable de la dégradation. La première conséquence importante observée dès la filière 130 nm ($T_{ox} = 2$ nm, V_{DD} = 1,2 V) est la disparition de la forme habituelle en cloche de I_{SUB} dans l'intervalle de polarisation. Par conséquent le pic de courant substrat n'apparaît plus pour les tensions intermédiaires ($V_{GS} \cong V_{DS}/2$) mais pour la région de polarisation où $V_{GS} = V_{DS}$, c'est-à-dire en allant vers un régime du transistor plus faiblement saturé. La deuxième conséquence importante est la persistance de la dégradation CHC pour des tensions de drain faibles, c'est à dire sous la barre en seuil d'énergie de 3,7 eV, ce qui contredit le principe imposé par le LEM [HU 85]. Ce point important sera abordé dans la section suivante qui décrit les modèles standards utilisés pour l'extrapolation de la durée de vie notamment avec le modèle de Guérin développé pour les porteurs à plus faible énergie.

Avec la réduction de T_{ox} , la dégradation CHC du transistor NMOS a subi la disparition des défauts dans le volume de l'oxyde de grille ΔN_{OT} qui avait lieu pour $V_{GS} = V_{DS}$ dans les filières antérieures, au profit des états d'interface accepteurs (ΔN_{TT} -) pour la condition pire cas qui est transférée dorénavant aux hautes tensions V_{GS} . La dégradation CHC du NMOS peut donc être décrite comme pour le pire-cas des générations précédentes (Figure 2.23a) mais avec une proportion de l'extension de la zone des défauts qui augmente avec la réduction de L_G.



Figure 2.23: Mécanisme de dégradation par porteurs chauds dans un transistor NMOS linéaire

Paradoxalement, la fiabilité CHC des transistors PMOS a subi plus de modifications pour cette même condition de polarisation pire-cas ($V_{GS} = V_{DS}$) avec la disparition du piégeage d'électrons ($|V_{TH}| \le |V_{GS}| \le |V_{DS}/5|$). Ceci est expliqué par les trous qui génèrent plus d'états d'interface que les électrons à travers le canal, ainsi que pour remplir les états lents du diélectrique mince (*slow traps*) (Figure 2.23). Il en résulte que la dégradation du PMOS se décompose comme pour la dégradation NBTI par la somme de défauts permanents (ΔN_{TT} Don) avec des défauts dits recouvrables ou états lents, qui se chargent avec le champ vertical (ΔN_{SS} o/+) et se déchargent (ΔN_{SS} +/o) avec la relaxation de la contrainte électrique [**BRA 09**].

Ces défauts CHC entrainent des réductions des performances du transistor PMOS en gains g_m , G_d et en courant $|I_{DSp}|$ avec $|\Delta V_{TH}| > 0$, qui se cumulent à la dégradation du NMOS, ce qui accélère la dégradation en dynamique des cellules CMOS [**BRA 09**]. Ainsi la dégradation par porteurs chauds du PMOS peut être considérée comme la somme d'une dégradation NBTI étendue à partir de la source et d'une dégradation CHC étendue à partir de la zone grille-drain comme le montre la Figure 2.23 [**BRA 09**].

2.4.2.2 Modèle de défaillance associé au porteurs chauds

Plusieurs modélisations des mécanismes de dégradations porteurs chauds des transistors MOSFETs existent dans la littérature. Nous aborderons les principaux modèles associés à la durée de vie parfois utiles par leur simplicité d'utilisation, en regard des mécanismes prédominants quand nous désirerons comparer la sensibilité de différents nœuds technologiques aux différents critères de fiabilité.

2.4.2.2.1 Modèle de Hu

Le modèle de Hu est fondé sur le mécanisme de l'électron chanceux dans lequel le porteur du canal possède assez d'énergie pour créer des paires-électrons trous. D'après cette interprétation, seuls les états d'interfaces sont responsables de la dégradation par porteurs chauds, les pièges dans le volume ne sont pas pris en compte.

Les états d'interface sont générés par des électrons énergétiques du canal qui sont assez "chanceux" pour ne pas perdre toute leur énergie cinétique par collisions sur les impuretés dopantes, par diffusion électron-phonons atteignant un seuil d'énergie de $\phi_{i,e}$ = 1.3 eV pour enclencher l'ionisation par impact. Les défauts sont modélisés par la relation :

$$\Delta N_{IT} = C_2 \cdot \left[\frac{I_{DS}}{W_{EFF}} \cdot e^{\left(\frac{-\Phi_{IT}}{q \cdot \lambda_e \cdot E_{LAT}} \right)} \cdot t \right]^n$$
(2.28)

où C_2 est une constante d'ajustement, W_{EFF} la largeur effective du transistor, Φ_{TT} l'énergie nécessaire pour créer les états d'interface, t le temps et n (0,4 à 0,5) le coefficient de dérive extrait des cinétiques temporelles considérées pour le paramètre transistor dégradé.

A partir, des équations (2.28) et (2.29), la durée de vie ou TTF est définie par rapport à un critère limite de réduction d'un paramètre considéré lié à la génération des états d'interface. Il est alors possible d'obtenir pour un transistor NMOS, une relation liant le temps à la défaillance aux courants de drain et de substrat représentatifs de l'accélération du phénomène porteurs chauds comme :

$$TTF_{\Delta N_{IT}} = C_3 \cdot \left[\frac{I_{DS}}{W_{EFF}} \cdot\right]^{-1} \cdot \left[\frac{I_{SUB}}{I_{SD}} \cdot\right]^{-m}$$
(2.29)

où m = 2,9 classiquement **[HU 85]** et m = 2,5 **[CHU 90]** pour des technologies plus récentes. Le modèle de Hu présente l'avantage de prendre en compte les dérives paramétriques en courant substrat et en courant de drain mais l'inconvénient d'être difficilement exploitable pour un utilisateur de composant qui n'a pas accès à ces valeurs de courants lorsque celui fonctionne.

2.4.2.2.2 Modèle de Takeda

Le modèle de Takeda considère également que la dégradation est provoquée par les états d'interface. Cependant, contrairement à la représentation de Hu, la dépendance est exprimée directement en fonction du courant substrat, et donc du taux d'ionisation mesuré à travers le courant de trous. La modélisation est elle aussi fondée sur le suivi de la dégradation d'un paramètre P de fonctionnement du transistor dans le temps. Le paramètre peut être aussi bien la tension de seuil que la transconductance :

$$\Delta P = A \cdot t^n \tag{2.30}$$

où A et n sont des variables empiriques pouvant être extraites pour chaque génération technologique dépendantes respectivement, de la relation entre I_{SUB} et V_{DS} et des cinétiques temporelles du paramètre P dégradé, utilisé pour l'extraction de TTF. La variable A présente une dépendance en V_{DS} et plus faiblement en V_{GS} liée à la dépendance simplifiée du maximum de courant substrat par :

$$A \propto e^{\left(-\frac{v}{V_{DS}}\right)} \tag{2.31}$$

où υ est une constante dépendante de la technologie ce qui donne $I_{SUB}^{max} \propto \exp(-\beta/V_{DS}) \propto A^{\beta/\upsilon}$ En combinant les équations (2.30) et (2.31) ainsi qu'en fixant un critère de défaillance sur un paramètre considéré, le temps à la défaillance TTF s'exprime en fonction de la tension drain source V_{DS} . De cette façon, la relation en TTF peut être exprimée par **[TAK 83]**:

$$TTF_p \propto e^{\left(\frac{\upsilon}{n \cdot V_{DS}}\right)} \propto (I_{SUB}^{max})^{-1} \text{ avec } I = \frac{-1}{n} \cdot \left(\frac{\upsilon}{\beta}\right)$$
 (2.32)

Ce modèle présente l'avantage d'être dépendant de la tension V_{DS} qui est une grandeur connue de l'utilisateur pour l'extrapolation à la tension d'alimentation V_{DD} . Néanmoins, le modèle comporte des variables très dépendantes de la technologie, de ce fait le modèle est difficilement extrapolable à d'autres nœuds technologiques sans connaitre la valeur des constantes. De plus, le modèle a été validé au milieu des années 80 sur des technologies 1 µm – 0,5 µm et ne semble plus être adapté aux nouveaux nœuds fortement submicroniques [GUE 08]. C'est pourquoi, le modèle de Hu est couramment utilisé dans l'industrie avec une prise en compte plus fine de l'effet du champ latéral proportionnel à I_{SUB}/I_{DS}.

2.4.2.2.3 Modèle de Guérin

Tout comme Takeda, le modèle de Guérin considère les états d'interface comme étant à l'origine de la dégradation par porteurs chauds pour la partie associée à la dégradation permanente **[GUE 08]**. Cependant, ce modèle va plus loin dans le sens où il généralise la dépendance en énergie et non plus en champ latéral, de la dégradation CHC à travers les différents mécanismes de rupture des liaisons Si - H à l'interface **[GUE 09]**. Il souligne en effet qu'avec la réduction de V_{DD} et des conditions de stress CHC, les mécanismes à l'origine de la dégradation ne sont plus les mêmes par rapport aux technologies antérieures, spécifiquement pour les porteurs de plus faibles énergies. En effet pour les conditions $V_{GS} \ge V_{DS}$ les électrons canal continuent à dépassiver les liaisons Si-H à l'interface à partir d'un certain seuil en densité de courant, et donc pour un nombre plus important de porteurs du canal. De ce fait, les mécanismes s'opèrent en 3 modes de dégradation **[GUE 09]**, **[BRA 09]** en plus de la rupture directe de la liaison pour les porteurs à fortes énergies.

- (a) Mode 1 : Le premier mode appelé SVE (Single Vibrational Excitation) apparait quand le transistor est polarisé en régime saturé, c'est-à-dire à plus faible courant (V_{GS} < V_{DS}) pour des porteurs de fortes énergies. Les électrons chauds excitent et font vibrer les liaisons Si-H comme une balance entre l'excitation et la désexcitation des liaisons aux abords de l'interface dans la zone saturée. Ce mode est donc relié à une excitation vibrationnelle multi niveaux de la liaison Si-H en un seul saut vibrationnel unique.
- (b) Mode 2 : Le deuxième mode appelé EES (Electron Electron Scattering) correspond aux porteurs d'énergie plus faible (intermédiaire) pour un régime de polarisation situé entre la forte et la faible saturation. Les électrons peuvent promouvoir collectivement certains d'entre eux à des énergies plus élevées par interaction électron-électron. Ce gain additionnel en énergie à celui fourni par le champ va permettre d'exciter la résonance des liaisons Si-H jusqu'à leur rupture.
- (c) Mode 3 : le troisième mode appelé MVE (Multiple Vibrational Excitation) intervient pour les porteurs de trop faible énergie pour exciter la résonance de la liaison Si - H. Ce régime correspond à la polarisation entre le régime faiblement saturé à linéaire (V_{GS} >> V_{DS}) où les porteurs incidents du canal contribuent à l'excitation multi vibrationnelle des liaisons Si - H avant désexcitation par interaction phonons jusqu'à la désorption des atomes de H.

La dégradation totale par porteur chauds est définie comme étant la somme des dégradations dans les différents modes. Ceci permet de balayer toutes les conditions de polarisations et les trois mécanismes de dégradation qui prendront tout leur poids suivant les différentes filières technologiques en L_G , V_{DD} , T_{ox} . Le point important du modèle est basé sur le fait que les 3 modes ont lieu en parallèle et que chacun des modes est dominant dans sa gamme d'énergie [GUE 09]. La dégradation et durée de vie totale est donc obtenue en sommant les 3 contributions :

$$\frac{1}{TTF_{HCI}} = \frac{1}{TTF_{SVE}} + \frac{1}{TTF_{EES}} + \frac{1}{TTF_{MVE}}$$
(2.33)

avec :

$$TTF_{SVE} = K_1 \cdot \left[\frac{I_{DS}}{W_{EFF}} \cdot\right]^{-a_1} \cdot \left[\frac{I_{SUB}}{I_{SD}} \cdot\right]^{-m}$$

$$TTF_{EES} = K_2 \cdot \left[\frac{I_{DS}}{W_{EFF}} \cdot\right]^{-a_2} \cdot \left[\frac{I_{SUB}}{I_{SD}} \cdot\right]^{-m}$$

$$TTF_{MVE} = K_3 \cdot V_{DS} \frac{-a_3}{2} \cdot \left[\frac{I_{SUB}}{I_{SD}} \cdot\right]^{-a_2} \cdot e^{\left(\frac{E_A}{k \cdot T}\right)}$$

$$(2.34)$$

où K_1 , K_2 , K_3 sont des constantes d'ajustement, I_{DS} et I_{SUB} sont les courants drain-source et substrat, V_{DS} est la tension drain-source (V), W représente la largeur du transistor (µm) et E_A l'activation en température du mécanisme du mode MVE (eV). Il est clair qu'avec (2.33), un mécanisme de dégradation dominant prendra tout son poids statistique en règle de Matthiessen, et deviendra la seule contribution majeure pour une technologie donnée.

Les exposants a_1 , a_2 , a_3 , m et E_A ont été déterminés expérimentalement par **[GUE 09]** et sont donnés dans le Tableau 2.5 pour les transistors NMOS et les PMOS. Il est à noter que l'activation positive en température $E_A = 0,26$ eV correspond à la diffusion des atomes d'hydrogène. De plus on observe pour les modes 2 et 3 une réduction nette de la durée de vie avec la température ce qui est un résultat opposé au comportement CHC classique en raison de son impact sur la fonction de distribution en énergie des porteurs.

Dispositif	a ₁	a_2	a ₃	m (S.D)	E _A (eV)
nMOS	2.7	1	2.5	17	0.26
pMOS	1.3	1	3	12	0.26

Tableau 2.5: Paramètres du modèle de Guérin pour une technologie 65 nm

Le modèle de Guérin présente l'avantage d'être fondé sur la physique des défauts d'interface à l'origine des mécanismes porteurs chauds en prenant en compte les 3 modes d'excitation de la liaison Si-H. Toutefois, comme pour le modèle de Hu, le modèle de Guérin nécessite de connaitre les courants I_{SUB} et I_{DS} , ce qui est difficile d'obtenir pour un utilisateur de composants commerciaux.

2.4.2.2.4 Modèle de prédiction de technologie (PTM HCI)

Le modèle Predictive Technology Model (PTM) est issu de la modélisation BSIM 3V3 du transistor MOS. Il a été développé par l'équipe de modélisation et d'intégration nanoscopique de l'université d'Arizona [PTM 06a]. Il modélise la variation de la tension de seuil du transistor dans le temps (ΔV_{TH}) en fonction des conditions d'utilisation et de la technologie. Il suffit alors d'établir un critère limite de dérive pour extraire le temps à la défaillance. Classiquement, un critère de variation de 10 % est retenu par les fabricants (ou 50 mV pour les technologies antérieures au 0,5 µm). La dérive de la tension de seuil est donnée par la relation suivante :

$$\Delta V_{TH} = \frac{q}{c_{OX} \cdot W_{EFF} \cdot L_{EFF}} \cdot K \sqrt{C_{OX} (V_{GS} - V_{TH0})} \cdot e^{\frac{E_{OX}}{E_0}} \cdot e^{\frac{-\varphi_{IT}}{q \cdot \lambda \cdot E_{LAT}}} \cdot e^{\frac{-E_A}{k} \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right)} \cdot t^n \quad (2.35)$$

où K, E_0 , λ , ϕ_{TT} et n sont des constantes, V_{TH0} est la tension de seuil avant vieillissement (V), C_{OX} et E_{OX} représentent respectivement la capacité surfacique (F.cm⁻²) et le champ vertical à travers l'oxyde de grille (V/cm). Ces derniers paramètres sont définis par les équations (2.36) et (2.37) :

$$C_{OX} = \frac{\varepsilon_0 \cdot \varepsilon_{0X}}{T_{OX}} \tag{2.36}$$

$$E_{OX} = \frac{V_{GS} \cdot V_{TH}}{T_{OX}} \tag{2.37}$$

où E_{LAT} représente le champ électrique latéral donné par la relation (1.3) et V_{DSAT} correspond à la tension de drain en saturation définie selon la modélisation BSIM3 par (2.38) :

$$V_{Dsat} = \frac{\left(V_{GS} - V_{TH0} + 2 \cdot \frac{k \cdot T}{q}\right) \cdot E_{SAT} \cdot L_{EFF}}{\left(V_{GS} - V_{TH0} + 2 \cdot \frac{k \cdot T}{q}\right) + A_{BULK} \cdot E_{SAT} \cdot L_{EFF}}$$
(2.38)

avec T la température (K), A_{BULK} un facteur lié à l'effet body incluant une dépendance avec la tension substrat-source V_{BS} et le dopage (et des coefficients géométriques), E_{SAT} le champ de saturation et enfin L_{EFF} la longueur de grille effective. l_0 est la longueur caractéristique en régime de saturation exprimée par la relation (1.5).

Dans le cadre de cette étude, le *modèle PTM HCI* a été ajusté avec les points expérimentaux issus des essais accélérés HCI de **[KOI 04]** pour les oxydes SiO2 et **[BRA 11]** pour les oxydes nitrurés SiON. Les résultats sur les oxydes en silice ont été obtenus sur des transistors NMOS de génération 180 nm (Figure 2.24) alors que ceux des oxydes nitrurés l'ont été sur des transistors NMOS de génération 90 nm **[BRA 11]**. Les paramètres extraits pour ces modèles sont résumés dans le Tableau 2.6.



Figure 2.24: Ajustement du modèle PTM HCI sur les points expérimentaux de **[KOI 04]**, transistors NMOS 116 nm, d'épaisseur électrique EOT 2,6 nm (SiO₂) aux conditions CHC.



Figure 2.25: Ajustement du modèle PTM HCI sur les points expérimentaux de **[BRA 11]**, transistors NMOS 90 nm, d'épaisseur électrique EOT 1,6 nm (SiON) aux conditions CHC.

Oxyde	Paramètres PTM après ajustement								
	K	\mathbf{E}_{0}	A _{BULK}	φ_{IT}	λ	n	E _{SAT}	E _A	ε _{ox}
	$(nm/C^{0,5})$	(V/nm)		(eV)	(nm)	(S.D)	(V/nm)	(eV)	
SiO ₂	1,6.10 ¹¹	0,8	0,005	3,7	7,8	0,46	0,011	0,2	3,9
SiON	$1,5.10^{7}$	0,11	0,005	3,7	7,8	0,43	0,011	0,2	4,3

Tableau 2.6: Constantes déterminées pour le modèle PTM HCI

Le modèle PTM HCI a le désavantage d'être semi-empirique dans le sens où il ne repose pas sur les mécanismes sous-jacents de l'injection par porteurs chauds. Toutefois, en ajustant les données expérimentales, il présente l'avantage d'être facilement extrapolable à différentes générations technologiques et à différentes conditions d'utilisations. C'est le modèle qui sera retenu par la suite pour simuler le mécanisme de porteurs chauds sur les composants fortement submicroniques.

2.4.3 L'instabilité de la tension de seuil en température (BTI)

L'instabilité de la tension de seuil en température ou BTI (*Bias Temperature Instability*) est un mécanisme d'usure du transistor MOS qui apparait à haute température lorsque celui-ci est polarisé en mode d'inversion. Historiquement, la dégradation est apparue dans les années 60 due à la contamination des diélectriques de grille par des ions mobiles dans les transistors PMOS [**DEAL 67**] sous tension de grille négative (NBTI). Avec l'amincissement de l'isolant, ce mode de dégradation est devenu très sévère dans les dernières filières CMOS malgré l'utilisation de diélectriques de grille nitrurés (NO ; NO₂) qui permettent la réduction de la génération des états d'interface. Il est en effet plus délicat d'optimiser les procédés de fabrication de la grille dans les filières les plus récentes ($L_G < 0,13 \mu m$, $T_{ox} < 2 nm$) qui utilisent les procédés de nitruration, dû à la concentration d'azote [HUA 05], en raison du piégeage de trous plus important dans ces technologies [**RIB 05**], [**DEN 04a**].

Le développement des diélectriques Hi-K a vu également ce mode d'instabilité de V_{TH} même pour les tensions positives dans le transistor NMOS **[CAR 03]**, **[KER 03]**, où l'on parle alors de mode PBTI. Généralement le BTI est examiné pour le PMOS étant donné la très forte accélération de la dégradation avec la température, sans courant I_{DS} et sous faible champ vertical, tandis que le PBTI dans le NMOS est moins sévère, car relatif au remplissage de pièges d'oxyde préexistants **[DEN 04a]**.

Le mécanisme de dégradation NBTI a été historiquement décomposé en deux parties [JEP 77]: la première, dominante à bas champ vertical est limitée par le phénomène de diffusion, et la seconde, apparaissant à plus fort champ, est contrôlée par l'injection par effet tunnel des porteurs et le piégeage de trous. *Jeppson* a été le premier à proposer que le NBTI entrainait une libération de l'atome d'hydrogène à l'interface Si/SiO₂. Ce modèle est à l'origine du modèle dit de « Réaction - Diffusion » (R-D) qui sera présenté section 2.4.3.1. Ce mode de dégradation génère des états d'interface N_{IT}(t) chargés positivement ainsi que des pièges dans le volume de l'oxyde de grille $N_{OT}^{+}(t)$ dû au piégeage de trous. Ceci conduit à une réduction du courant $|I_{DSp}|$ (équation (2.39)) et un décalage important de la tension de seuil ($\Delta V_{TH} < 0$) cumulé à la réduction de la mobilité au cours du temps, tel qu'en régime linéaire, nous avons:

$$I_{Dlin} = C_{OX} \cdot \frac{W_G}{L_G} \cdot \frac{\mu_0}{1 + \alpha \cdot N_{IT}(t)} \cdot \left[V_{GS} - \left(V_{TH0} + \Delta V_{TH}(t) - \frac{V_{DS}}{2} \right) \right] \cdot V_{DS}$$
(2.39)

Dans le cas d'un transistor PMOS, (Figure 2.26) les travaux de Huard **[HUA 05]**, ont montré l'augmentation du piégeage de trous lorsque la concentration d'azote augmente à l'interface substrat-isolant avec le procédé d'incorporation thermique rapide (RTN). Ceci souligne que pour un niveau de ΔN_{TT} généré, la dérive de la tension de seuil est plus forte lorsque la concentration en azote augmente. C'est pourquoi l'optimisation de la nitruration en concentration et en température a été nécessaire en passant à la nitruration plasma (DPN) pour décaler le maximum d'azote vers l'interface grille-isolant, ce qui a permis d'obtenir une meilleure résistance à la dégradation NBTI pour le transistor PMOS **[HUA 05]** par rapport aux oxydes de grille SiO₂ purs.



Figure 2.26: Conditions BTI pour un PMOS

Figure 2.27: Conditions BTI pour un NMOS

Comme la dégradation NBTI est la plus sévère par rapport au PBTI (NMOS), le mécanisme de l'instabilité de la tension de seuil est systématiquement considéré par les fabricants sur les transistors PMOS à l'aide de test effectués entre 125°C et 150°C.

2.4.3.1 Physique de la défaillance pour les oxydes en silice et nitruré

2.4.3.1.1 Etats d'interface NIT

Le formalisme de réaction-diffusion (R-D) a traditionnellement été proposé pour décrire la génération de défauts sous contraints NBTI **[JEP 77]**. Le modèle R-D s'appuie sur la diffusion d'espèces hydrogénées à partir de la rupture des liaisons SiH de l'interface isolant-substrat. Les espèces hydrogénées proviennent des différentes étapes de fabrication réalisées pour passiver les liaisons pendantes du substrat. Il est admis que les espèces mises en jeu dans la dégradation NBTI sont les protons H^+ , les atomes d'hydrogènes H ainsi que le dihydrogène H_2 **[ALAM 05]**.

La modélisation R-D fut tout d'abord améliorée par Ogawa *et al.* en 1995 **[OGA 95]** puis fut ensuite reprise par Alam *et al.* en 2005 pour y introduire l'influence de la température sur le coefficient de dissociation **[ALAM 05]**. L'intérêt de cette approche consiste premièrement à comprendre le type d'espèces mises en jeu lors de leur diffusion et deuxièmement, à extraire le coefficient de dépendance temporelle du mécanisme qui sera ensuite utilisé pour extrapoler le temps à la défaillance sur des durées plus longues.



Figure 2.28: a) Schéma de la dissociation des liaisons Si-H par rapport à l'interface δ **[ALA 05]**, b) Représentation de la densité de proton H⁺ à l'interface et dans le volume de l'oxyde **[ALA 07]**, c) Représentation de la densité d'atome H⁺ à l'interface et dans le volume de l'oxyde **[ALA 07]**

Tout d'abord, en régime d'inversion, les porteurs minoritaires s'accumulent à l'interface oxyde-substrat. Sous l'effet du champ vertical, ces porteurs sont accélérés et peuvent casser les liaisons Si-H à l'interface selon les réactions électrochimiques:

$$Si_3 \equiv Si - H + [t^+] \rightarrow Si_3 \equiv Si^\circ + H^+$$
 PMOS NBTI (2.40)

$$Si_3 \equiv Si - H + [e^-] \rightarrow Si_3 \equiv Si^\circ + H^-$$
 NMOS PBTI (2.41)

La formation de pièges à l'interface dans le temps est proportionnelle à la différence entre le nombre initial de liaison Si-H potentiellement dissociables N_0 et la densité d'espèces hydrogénées N_H pouvant être recombinées avec les liaisons pendantes. Cette dépendance est modélisée par la relation dite de réaction :

$$\frac{dN_{IT}}{dt} = k_F \cdot (N_0 + N_{IT}) - k_R \cdot N_H \cdot N_{IT} \qquad (x = 0)$$
(2.42)

où k_F et k_R sont respectivement le taux de formation d'états d'interface (sens direct ou Forward) et le taux de recombinaison entre les espèces Si et H (sens reverse).

Ensuite, sous l'influence du champ vertical, les espèces hydrogénées peuvent soit diffuser soit interagir entre elles dans le volume de l'oxyde selon les réactions chimiques:

$$H^+ + e^- \longrightarrow H^0 \tag{2.43}$$

$$H^0 + H^0 \longrightarrow H_2 \tag{2.44}$$

La diffusion d'espèces hydrogénées N_H (H+, H et H₂) à partir de l'interface dans le volume de l'oxyde est modélisée par la relation dite de diffusion :

$$\frac{dN_{IT}}{dt} = -D_H \cdot \frac{dN_H}{dx} + N_H \cdot \mu_H \cdot E_{OX} + \frac{\delta}{2} \cdot \frac{dN_H}{dt} \qquad (0 < x < \delta)$$
(2.45)

où D_H (cm².s⁻¹) et μ_H (cm².V⁻¹.s⁻¹) sont respectivement le coefficient de diffusion et la mobilité d'espèces hydrogénées.

La solution de l'équation de diffusion diffère selon le type d'espèces impliquées dans le mécanisme NBTI et donc suivant leur dépendance temporelle. Si les espèces mises en jeu sont les protons H⁺, alors la solution de l'équation de diffusion sera de la forme :

$$N_{IT}(t) \propto \sqrt{\frac{k_F \cdot N_0}{k_R}} \cdot (D_H \cdot t)^{\frac{1}{2}} \qquad (n = 1/2)$$
 (2.46)

en revanche, si ce sont les molécules de dihydrogène H₂, la solution sera :

$$N_{IT}(t) \propto \left[\frac{k_F \cdot N_0}{2 \cdot k_R}\right]^{\frac{2}{3}} \cdot (D_H \cdot t)^{\frac{1}{6}} \qquad (n = 1/6)$$
 (2.47)

enfin si ce sont les atomes d'hydrogène, la variation temporelle d'état d'interface sera exprimée par (2.48) :

$$N_{IT}(t) \propto \sqrt{\frac{k_F \cdot N_0}{k_R}} \cdot (D_H \cdot t)^{\frac{1}{4}} \qquad (n = 1/4)$$
 (2.48)

Expérimentalement, la formation d'états d'interface dans le temps est donnée dans la littérature comme suivant une loi de puissance dont le coefficient est entre 0,16 - 0,25 **[ALA 07] [AON 05]**. La valeur classique pour le mécanisme NBTI étant n = 1/4 pour les espèces neutres **[TSU 05]**. Toutefois, il est courant de trouver dans la littérature des résultats où la dépendance temporelle est exprimée par un coefficient n égal à 0,16 **[ALA 07]**. A première vue, cette différence pourrait signifier que le mécanisme devient moins critique au cours du temps pour les technologies plus récentes et que leur fiabilité s'améliore. Cependant il n'en n'est rien. La différence de coefficient s'explique par l'influence de la relaxation directement liée au délai latent entre le stress et la mesure de la tension de seuil. Il a été démontré en effet que plus le délai de mesure est important, plus l'exposant n augmente **[CHA 04]**, **[ALA 07]** (Figure 2.29).

Ce point important a fait l'objet de nombreux développements pour améliorer les techniques de mesures, grâce notamment aux travaux de Denais et al. **[DEN 04b]** qui ont été les pionniers pour mettre au point une technique rapide de mesure à la volée, ou "*On The Fly*" (OTF), sans relâchement des phases de contraintes NBTI. Ces techniques avancées propres à l'emploi et le test des diélectriques ultra-minces ($T_{ox} < 3nm$), ne nécessitent plus d'interrompre le vieillissement pour pouvoir mesurer I_{DSlin} (V_{TH} , G_m) extraits simultanément où la dépendance temporelle devient proche de n = 0,16 - 0,19 sur une double échelle logarithmique comme le montre la Figure 2.29.



Figure 2.29: Influence du temps de mesure sur le coefficient de dépendance temporel [ALA 07]

D'autre part, l'avantage des techniques de mesures OTF a été de démontrer l'influence des charges d'oxyde remplis par l'injection de trous, comme des défauts lents (N_{ss}) dit relaxables. Ces défauts sont situés dans des oxydes de grille à une distance tunnel du substrat (et de la grille) et ceci, spécifiquement dans les oxydes de grille nitrurés, où le piégeage de trous est plus important [**DEN 04**], [**HUA 05**]. Ceci nous conduit à présenter ces types de défauts dans la section suivante.

2.4.3.1.2 Pièges dans le volume NOT+

Initialement, de nombreux défauts existent dans le volume de l'oxyde de grille dépendant intrinsèquement de la qualité des procédés d'élaboration et extrinsèquement des mécanismes de dégradation. Il est généralement admis que plus les oxydes sont fins, plus le nombre de pièges en volume devient faible **[POB 10]**. Ceci signifie que les oxydes de type SiO₂ et SiON des nouvelles générations seraient moins susceptibles de comporter des pièges. D'autant plus qu'à fort champ électrique comme à plus faible champ, les mécanismes de génération de pièges d'oxyde sont limités par les effets de dépiégeage par le champ, comme également par l'effet Poole-Frenkel.

Toutefois, les oxydes à forte permittivité sont par nature plus épais et comportent intrinsèquement plus de pièges et d'états lents pouvant participer aux mécanismes BTI **[STA 06]**, **[GAR 09]** ce qui rend l'analyse de l'effet de ces défauts essentielle pour déterminer la fiabilité d'un empilement de grille.



Figure 2.30: Dégradation de la tension de seuil d'un transistor PMOS soumis à une contrainte de type NBTI **[STA 05]**



Figure 2.31: Phase de dégradation et de relaxation de la tension de seuil pour une contrainte de type NBTI **[HUA 07]**

Une des conséquences de l'influence de la mesure et de l'interruption des contraintes électriques sous conditions de stress NBTI est l'apparition des phénomènes de relaxation comme nous l'avons discuté dans la section précédente. En absence de contrainte électrique entre la grille et le substrat, les trous pièges sur des états lents (ou border traps) peuvent alors se dépiéger vers le canal (la grille). La conséquence électrique du dépiègeage de charge est une diminution de la tension de seuil (Figure 2.31). La relaxation de V_{TH} atteint un niveau correspondant aux défauts permanents attribués aux états d'interface rapides [HUA 07]. Ceci signifie que la dégradation NBTI est composée d'une partie de défauts relaxables attribués aux centres E'. Les états lents sont associés aux lacunes d'oxygène V_O ou atomes interstitiels d'oxygène [KAR 99]. Les centres E' ($O_3 \equiv Si^{\bullet}$) impliquent deux atomes de silicium avec un électron non apparié et peuvent se dissocier en centre E'_{δ} (faible relaxation) et E'_{γ} (forte relaxation) [NIC 02].

En condition NBTI, d'autres pièges peuvent être générés dans le volume de l'oxyde en plus du remplissage de pièges pré-existants. Les espèces hydrogénées issues de la dissociation des liaisons Si-H à l'interface diffusent dans l'oxyde en direction de la grille (phase de diffusion). Sur leur trajet, ces espèces peuvent créer des défauts dans le volume en entrant en collision avec les liaisons Si-O dans le cas de la silice ou les liaisons Si-N pour les oxydes nitrurés. Les relations électrochimiques de ces réactions sont les suivantes **[KIM 00]** :

$$Si_3 \equiv 0 + H^+ \longrightarrow Si_3 \equiv OH^+$$
 oxydes SiO₂ (2.49)

$$Si_3 \equiv N + H^+ \rightarrow Si_3 \equiv NH^+$$
 oxydes nitrurés (2.50)

Ces défauts notés N_{OT}^+ se comportent comme des pièges pour les trous du canal. La conséquence électrique du piégeage est une augmentation de la tension de seuil du transistor (Figure 2.30 et Figure 2.31). Le pire cas NBTI correspond donc à la condition DC au cours de laquelle le transistor est contraint de manière permanente sans influence de la relaxation.

2.4.3.2 Modèle de prédiction de technologie (PTM BTI)

Le modèle PTM appliqué à la dégradation NBTI permet de calculer la variation de la tension de seuil du transistor dans le temps ($\Delta V_{TH}(t)$) en fonction des conditions d'utilisation et de la température. En fixant un critère limite de dérive il est possible d'extraire le temps à la défaillance. Comme pour la dégradation sous injections de porteurs chauds, un critère de dérive (défaillance) de 10% est retenu par les fabricants ou 50 mV pour les technologies antérieures au nœud 0,5 µm. Ceci permet de garantir un niveau de sécurité assez conservatif pour tester les technologies.

Le modèle PTM BTI se décline en deux parties **[PTM 06b]**. Le premier permet d'étudier la dégradation et la relaxation de la tension de seuil pour quelques cycles afin d'ajuster les paramètres du modèle sur des points expérimentaux. Le deuxième est en réalité une extension du premier, où la notion de rapport cyclique est introduite pour tenir compte des différentes phases où sont générés les défauts permanents et relaxables, comme décrit précédemment. Par conséquent, pour ce qui nous concerne, la dégradation peut être calculée par exemple sur un nombre de cycles représentatif d'une application aéronautique.

2.4.3.2.1 Prédiction PTM BTI cycle par cycle

Le modèle de prédiction PTM BTI cycle par cycle est applicable au transistor PMOS par une succession de cycles de période T, de rapport cyclique α , comme le montre la Figure 2.32 **[PTM 06b]**. Ceci permet d'enchainer de façon périodique les phases au cours desquelles auront lieu la dégradation NBTI et la phase de relaxation par relâchement de la contrainte électrique.



Figure 2.32: Conditions de fonctionnement du transistor PMOS : phase de contrainte et de relaxation [PTM 06b]

En considérant le cycle présenté Figure 2.32, la variation de la tension de seuil d'un transistor $\Delta V_{TH_{-C}}^{N}(t)$ au cours du N^{eme} cycle de contrainte (entre t_{N-1} et t_N), s'exprime par le modèle suivant :

$$\Delta V_{TH_c}^{N}(t) = \sqrt{K_V^2 \cdot (t - t_{N-1})^{0.5} + \Delta V_{TH_c}^{N}(t_{N-1}) + \delta_V}$$
(2.51)

avec :

$$\Delta V_{TH_{c}}^{N}(t_{N-1}) = \Delta V_{TH_{r}}^{N-1}(t_{N+1})$$

$$\Delta V_{TH_{c}}^{1}(t_{0}) = \delta_{V}$$

$$t_{0} = 0$$

(2.52)

où t représente le temps, $\Delta V_{TH_{L}C}^{N}(t_{N-1})$ la variation de tension de seuil au début du N^{ième} cycle de contrainte, δ_V une constante d'ajustement, $\Delta V_{TH_{L}C}^{-1}(t_0)$ la tension de seuil initiale. K_V est une variable dépendante de la technologie ainsi que du profil de mission donnée par :

$$K_{V} = A \cdot T_{OX} \cdot \sqrt{C_{OX} \left(V_{GS} - \Delta V_{TH_{c}}^{1}(t_{0}) \right)} \cdot e^{\frac{E_{OX}}{E_{0}}} \cdot \left[1 - \frac{V_{GS}}{a \left(V_{GS} - \Delta V_{TH_{c}}^{1}(t_{0}) \right)} \right] \cdot e^{\frac{-E_{A}}{k} \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}} \right)}$$
(2.53)

où A est une constante exprimée en V.nm⁻¹.C⁻¹, T_{OX} l'épaisseur de l'oxyde, C_{OX} la capacité d'oxyde, V_{GS} la tension grille-source, E_{OX} le champ électrique à travers l'oxyde de grille, E_0 le facteur d'accélération en champ, a une constante sans dimension, V_{DS} la tension drain-source, E_A l'énergie d'activation du mécanisme. Sur la phase de relaxation (entre t_N et t_{N+1}), la variation de la tension de seuil $\Delta V_{TH_LR}^{N}(t)$ au cours du N^{eme} cycle est donnée par les relations :

$$\Delta V_{TH_r}^{N}(t) = \left(\Delta V_{TH_r}^{N}(t_N) - \delta_V\right) \cdot \left[1 - \sqrt{\frac{\eta \cdot (t - t_N)}{t}}\right]$$
(2.54)

avec :

$$\Delta V_{TH_r}^{1}(t_1) = \Delta V_{TH_c}^{1}(t_1)$$
(2.55)

et où η est une constante.

Pour notre étude, le *modèle PTM BTI cycle par cycle* a été ajusté sur les points expérimentaux issus des essais accélérés NBTI de **[CHE02]**, effectués sur des transistors PMOS de génération 120 nm (Figure 2.33). Pour cela, une partie des constantes du modèle a été ajustée en donnant des résultats satisfaisant par rapport aux points expérimentaux. Les constantes extraites sont récapitulées dans le Tableau 2.7.



Figure 2.33: Ajustement du modèle PTM NBTI cycle par cycle sur les points expérimentaux de **[CHE02]**, transistors PMOS 120 nm, oxyde SiO₂ d'épaisseur électrique EOT = 1,3 nm, rapport cyclique 0,5

Paramètres	A	E ₀	a	η	E _A	δ _V	α
	(V.nm ⁻¹ .C ⁻¹)	(MV/cm)	(S.D)	(S.D)	(eV)	(mV)	(S.D)
Valeur	1,14.10-8	2	1,3	0,35	0,18	5.10-3	100

Tableau 2.7: Paramètres du modèle PTM NBTI cycle par cycle

Le modèle PTM cycle par cycle permet donc de calculer de manière assez précise la variation de tension de seuil pour les premiers cycles d'utilisation. Toutefois, son utilisation devient assez lourde lorsqu'il s'agit d'étudier le vieillissement d'un dispositif sur une longue durée. Pour cela le modèle PTM multi cycle se montre plus approprié comme nous allons le voir dans la section suivante.

2.4.3.2.2 Prédiction PTM multi-cycle longue durée

Le *modèle PTM multi-cycle* introduit la notion de rapport cyclique d'utilisation **[PTM 06b]**. De ce fait, il est plus représentatif d'une utilisation réelle du dispositif que le modèle *PTM cycle par cycle*. Dans ce cas, l'expression du décalage en tension de seuil est donnée au cours de N cycles de contrainte et relaxation par :

$$\Delta V_{TH}(T, N, \alpha) = K_V \cdot \left(\frac{\alpha}{100}\right)^n \cdot T^n \cdot \left[\frac{1 - \left(1 - \sqrt{\frac{\eta \cdot \left(1 - \left(\frac{\alpha}{100}\right)\right)}{N}}\right)^{2N}}{1 - \left(1 - \sqrt{\frac{\eta \cdot \left(1 - \left(\frac{\alpha}{100}\right)\right)}{N}}\right)^2}\right]^{(2.56)}$$

où K_v est donnée par l'équation (2.53), $\alpha/100$ est le rapport cyclique, T la période du cycle, n est le coefficient de dépendance temporelle du mécanisme NBTI, η est une constante, N est le nombre de cycles.

Les constantes du modèle dépendent de la nature de l'oxyde de grille. C'est pourquoi dans le cadre de nos travaux utilisés pour décrire la dégradation NBTI, un ajustement spécifique du modèle PTM multi cycle a été effectué sur des points expérimentaux de la littérature en fonction du type de diélectrique:

- pour la silice SiO₂ [ALA 05],
- pour les oxydes nitrurés SiON [CHA 05b],
- pour l'oxyde à haute permittivité HfSiON [ZHA 09].

Les points expérimentaux sur oxyde de grille SiO₂ pur **[ALA 05]** ont été obtenus sur des transistors PMOS de longueur $L_G = 0,12 \mu m$ vieillis de manière statique à T = 125°C sur une période des stress $t_s = 10^4$ s et d'EOT = 2,6 nm. Les tensions de contrainte sont respectivement égales à $-V_{GS} = 2,30$ V, 1,80 V et 2,08 V. Ces résultats représentent donc le cas de l'oxyde très mince de technologie de cœur standard. Les résultats sur oxyde de grille nitruré SiON **[CHA 05b]** sont obtenus pour des transistors PMOS d'entrée-sortie (IO) de longueur $L_G = 0,4 \mu m$ à oxyde de grille épais EOT = 6,8 nm et à plus forte température T = 150°C que l'exemple précédent. Les tensions de stress sont $-V_{GS} = 2,5$ V et 2,9 V appliquées pendant 3000 secondes.

Un exemple de modélisation du NBTI dans des diélectrique Hi-K+IL d'hafnium et d'oxyde nitruré (HfSiON) est effectué à partir des données de **[ZHA 09]**. Les transistors PMOS $L_G = 1 \mu m$ ont une épaisseur d'oxyde équivalente EOT = 2 nm. La température de test est la même que dans l'exemple précédent à 150°C alors que la tension est fixée à 2,2 V.

Un facteur de seuil en champ E_0 est obtenu égal à 4 MV/cm ainsi qu'une énergie d'activation $E_A = 0,11$ eV ont pu être déterminés grâce aux travaux de **[DEO 08]**. De ces résultats, on observe par exemple à t_s = 1000 s (Figure 2.34 et 2.35), pour des conditions de champ similaire ($V_{GS} = -2,5$ V et -2,3 V) que l'effet de la température est plus important avec une augmentation de $\Delta V_{TH} = 25$ mV (SiON) à 150°C pour $\Delta V_{TH} = 14$ mV (125°C, SiO₂) malgré l'emploi d'un oxyde de grille équivalent plus épais dans les entrée-sortie IO (6,8 nm) par rapport au cœur logique (2,6 nm). Ceci confirme avec les data de la littérature le plus fort impact de la température sous contrainte NBTI en accord avec le modèle PTM.





Figure 2.34: Ajustement du modèle PTM NBTI multi cycle sur les points expérimentaux de **[ALA 05]** pour des transistors PMOS 120 nm, d'oxyde SiO₂ épaisseur électrique EOT= 2,6 nm (SiO₂)

Figure 2.35: Ajustement du modèle PTM NBTI multi cycle sur les points expérimentaux de **[CHA05]** pour des transistors PMOS 0,4 µm, d'oxyde SiON épaisseur électrique EOT 6,8 nm.



Figure 2.36: Ajustement du modèle PTM NBTI multi cycle sur les points expérimentaux de **[ZHA 09]** pour des transistors PMOS 1 μ m, d'oxyde HfSiON épaisseur électrique EOT = 2 nm.

L'ajustement des différents paramètres a permis de calculer la dégradation de ΔV_{TH} en accord avec les cinétiques temporelles qui seront utilisées par la suite pour des durées plus longues. Ces paramètres sont regroupés pour les trois oxydes de grille dans le Tableau 2.8. Afin d'approcher le vieillissement statique à partir du modèle multi cycle, un seul cycle de rapport cyclique a été choisi en posant dans l'équation 2.54 le facteur $\alpha = 99,99\%$.

Paramètres	A (V/nm/C ^{0,5})	E ₀ (MV/cm)	a (S.D)	η (S.D)	E _A (eV)	n (S.D)	N (S.D)	α %
SiO ₂	5.10-6	2	1,3	0,35	0,13 [ALA 05]	0,25	1	99,99
SiON	6 ,2 .10 ⁻⁵	2	1,3	0,35	0,14 [TSU 05]	0,17	1	99,99
HfSiON	1,4.10-4	4 [DEO 08]	1,3	0,35	0,11 [DEO 08]	0,23	1	99,99

Tableau 2.8: Paramètres du modèle PTM NBTI multi cycle pour les oxydes SiO2, SiON et HfSiON

2.5 Conclusion

Une revue exhaustive des principaux modes de défaillance considérés comme les plus critiques au niveau du circuit intégré, a été présentée dans ce chapitre. Ces différents modes ont nécessité de décrire les mécanismes à l'origine des défaillances prépondérantes qui sont les points clefs dans la compréhension de la fiabilité en lien avec les enjeux front-end et back-end de la conception des circuits intégrés. Les modèles de dégradation de la littérature ont été décrits car ils permettent de distinguer le degré de sensibilité correspondant aux diverses filières technologiques présentées dans le chapitre 1. Ceci nous offre les connaissances de bases pour pouvoir analyser la fiabilité des dispositifs dans le but de développer une méthodologie d'analyses fonctionnelles pour les applications avioniques et aérospatiales.

Le premier élément de conclusion est que pour un même mécanisme, il existe plusieurs modèles de dégradations possibles suivant le niveau d'intégration du composant. La première étape est donc une bonne connaissance du système et des données technologiques pour choisir les modèles pertinents d'accélération de la défaillance en fonction des principaux paramètres d'accélération, tension, température cyclages. Ceci est notamment le cas pour le mécanisme de claquage de l'oxyde de grille, où en dessous de la génération 90 nm, le modèle classique dit en E n'est plus utilisable et doit être remplacé par un modèle en puissance de l'accélération en tension. Il se pose donc le problème du choix d'un modèle de défaillance, de sa pertinence et sa justesse par rapport à une technologie submicronique en fonction des principaux éléments connus de la technologie.

Le deuxième point est que pour un même mécanisme et même modèle, il existe plusieurs valeurs de paramètres selon la complexité des structures et la nature des matériaux utilisés. Nous sommes donc tributaires des données constructeurs pour établir les paramètres d'accélération tels que les énergies d'activation, les gammes de tensions en regard de la nature des matériaux quand ils sont précisés. Ceci implique qu'une bonne détermination de la fiabilité du composant nécessite d'une part de bien connaitre les éléments technologiques et d'autre part de bien maitriser tous les paramètres de dégradation associés. Pour cela, une synthèse des paramètres de dégradation des divers modes de défaillance et leurs conséquences a été effectuée pour les technologies fortement submicroniques, disponible dans le Tableau 2.9.

Nous avons vu dans ce chapitre que les défaillances pouvaient se séparer en deux principales familles (*i*) le claquage franc d'oxyde de grille et d'oxyde intermétallique, (*ii*) les mécanismes d'usure avec l'électromigration, les contraintes mécaniques et les injections de porteurs chauds. Pour cela les modèles d'extrapolation de durée de vie nous ont permis de quantifier de manière simple le niveau de fiabilité d'une technologie en fonction des dimensionnements; des conditions d'utilisation en courant, en tension et température d'utilisation par rapport à la température de test. Un critère de défaillance comme la dérive de la tension de seuil a été choisi pour la dégradation porteurs chauds ou BTI, nous permettant d'extraire un temps de durée de vie pour chacun de ces modes de dégradations.

Grâce à cette approche détaillée dans ce chapitre, nous pouvons maintenant évaluer la criticité des mécanismes vis à vis d'applications aéronautiques qui représentent de plus fortes contraintes en température, champs électriques, courants et cyclages, en les appliquant à ces nouvelles spécifications de profils opérationnels. Cette démarche est proposée dans le chapitre suivant au travers de 3 applications aéronautiques principales en mode d'accélération et aux conditions réelles d'utilisations.
	Défaillance					
Zone	Туре	Site	Mode	Mécanisme	Conditions critiques MOS	Type de défaillance
	Electromigration	Ligne de métallisations longues	Circuit ouvert dans une ligne	Création de lacunes dans une ligne	Fort courant Haute température	Défaillance déterministe (usure)
BEOL	Migration par contrainte mécanique	Via d'interconnexion	Circuit ouvert dans un via	Création de lacunes dans un via	Haute température Cycle thermiques	Défaillance déterministe (usure)
	Claquage d'oxyde intermétallique	Oxyde d'isolation interligne	Court circuit interligne	Création d'un chemin percolateur interligne	Différence de potentiel Haute température	Défaillance déterministe (usure)
	Claquage d'oxyde de grille	Oxyde de grille du transistor	Courant de fuite Court circuit grille canal	Création de pièges dans le volume de l'oxyde	$V_{GS} = V_{DD}$ Haute température	Défaillance aléatoire ou déterministe selon T _{OX}
FEOL	Porteurs chauds	Oxyde de grille du transistor	Dégradation de la tension de seuil et du courant de fonctionnement	Création d'états d'interface et de pièges dans l'oxyde	$V_{GS} = V_{DD}/2$ Basse Température $T_{OX} > 2 \text{ nm}$ $V_{GS} = V_{DS} = V_{DD}$ Haute Température $T_{OX} < 2 \text{ nm}$	Défaillance déterministe (usure)
	Instabilité de la tension de seuil en température	Oxyde de grille du transistor	Dégradation de la tension de seuil et du courant de fonctionnement	Cassure de liaisons Si-H et diffusion d'espèces H ⁰ , H ₂	$V_{GS} = V_{DD}, V_{DS} = 0$ Haute température	Défaillance déterministe (usure)

Tableau 2.9: Synthèse des principaux modes de défaillances des circuits intégrés

3 IMPACT DES MECANISMES DE DEFAILLANCE DU CIRCUIT INTEGRE SUR LA DUREE DE VIE DE COMPOSANTS UTILISES DANS DES APPLICATIONS AERONAUTIQUES

3.1 INTRODUCTION	113
3.2 PROFILS DE MISSION AERONAUTIQUES	113
3.2.1 Profil de Mission d'un calculateur	
3.2.1.1 Le calculateur de commande de vol en avionique civile	
3.2.1.2 Profil de mission électrique et thermique par phase	116
3.2.2 AUTRES PROFILS DE MISSION	117
3.2.3 Specifications en fiabilite	
3.2.3.1 Définition de la fiabilité d'un composant	
3.2.3.2 Spécifications de fiabilité pour les 3 applications	
3.3 APPLICATION DES MODELES DE DEGRADATION POUR TROIS APPLICATIONS AERONAUT	IQUES 122
3.3.1 LES MECANISMES DE DEFAILLANCE BEOL	
3.3.1.1 L'électromigration	
3.3.1.2 La dérive induite par des contraintes thermomécaniques (SIV)	130
3.3.1.3 Le claquage d'oxyde intermétallique	
3.3.2 LES MECANISMES DE DEFAILLANCE FEOL	146
3.3.2.1 Le claquage de l'oxyde de grille	147
3.3.2.2 Les porteurs chauds	151
3.3.2.3 L'instabilité de la tension de seuil en température (BTI)	159
3.3.3 COMPARAISON DES CRITERES DE DEFAILLANCES	164
3.4 CONCLUSION	

3.1 Introduction

Les mécanismes physiques susceptibles d'engendrer la défaillance d'un circuit intégré ont été présentés dans le chapitre précédent. Cependant, quel peut être l'impact de ces mécanismes sur la durée de vie d'un circuit intégré complet ? Et surtout comment s'assurer que le composant commercial soit fiable vis-à-vis de ces mêmes mécanismes? En toute rigueur, il n'est possible de répondre que si l'environnement et les spécifications d'utilisations sont définis au préalable. Or, il est bien entendu impossible, ou tout au moins très difficile, pour un fabricant de connaitre et d'anticiper les performances et la fiabilité du composant vendu. De ce fait, il n'existe pas de composant commercial vendu totalement *fiable*.

En revanche, un utilisateur peut *fiabiliser* un composant en disposant soit du retour d'expérience sur la durée de l'application et des contraintes réelles impliquées, soit des résultats des tests de vieillissement accélérés. En considérant que le retour d'expérience n'est possible que pour les technologies éprouvées les moins récentes, le seul outil d'évaluation demeure l'extrapolation sur la durée de l'application de données issues de vieillissements accélérés pour les technologies les plus avancées.

De cette manière, nous proposons dans ce chapitre de calculer la durée de vie de composants fortement submicroniques de générations 0,5 µm jusqu'au 45 nm, soumis à trois profils de mission caractéristiques d'applications aéronautiques. Une évaluation de la criticité par application et par technologie sera ensuite discutée à l'aide des facteurs d'accélération et du choix des facteurs aggravants. Enfin, des solutions de compromis seront proposées devant les résultats et le choix que cela nous impose pour ces applications aéronautiques à fortes contraintes opérationnelles. Ce chapitre fournira donc les bases pour l'analyse des essais accélérés (chapitre 4), ainsi que le développement d'une méthodologie spécifique aux applications aéronautiques qui sera présentée dans le chapitre 4.

3.2 Profils de mission aéronautiques

Les facteurs environnementaux impactant la fiabilité du composant sont nombreux en aéronautique. Il est donc nécessaire de distinguer à chaque niveau du composant, les facteurs d'utilisation les plus contraignants qui peuvent conduire à la défaillance. Classiquement, au niveau du report du composant sur le circuit imprimé, les aspects vibratoires, les chocs ainsi que les cycles thermiques sont les facteurs les plus critiques. Au niveau de la résine d'encapsulation, l'humidité et les cycles thermiques sont les contraintes prépondérantes. Enfin, au niveau du circuit intégré, seules les commandes électriques et l'environnement thermique accélèrent les mécanismes d'usure et de fatigue qui peuvent enclencher des défaillances franches ou progressives.

Les défaillances au niveau de l'assemblage et du boitier ne sont pas traitées dans ce manuscrit. De ce fait, seuls les profils de mission en température et en utilisations électriques seront pris en considération pour la description des profils de mission.

3.2.1 Profil de mission d'un calculateur

3.2.1.1 Le calculateur de commande de vol en avionique civile

Les commandes de vol sont les systèmes assurant le lien entre le manche de pilotage et les gouvernes aérodynamiques qui permettent de modifier la trajectoire de l'avion (Figure 3.1). Depuis 1984, les commandes de vol sont devenues informatisées et le traitement des données issues du manche et des différents capteurs de l'avion est réalisé à l'aide de calculateurs numériques.



Figure 3.1: Schéma d'un système de commande de vol électrique [AVI 11]

Le calculateur est devenu l'élément central des commandes de vol électriques. Généralement il est possible de distinguer les calculateurs dits primaires, utilisés pendant l'intégralité d'un vol (permettant de contrôler la profondeur, les ailerons, la direction, les aérofreins et les vérins de plan horizontal), et les calculateurs secondaires employés uniquement lors des phases de décollage et d'atterrissage.

La fiabilité des commandes de vol est garantie par la redondance des équipements composant le système. De cette manière, les calculateurs primaires et secondaires sont dupliqués en deux ou trois exemplaires selon la génération de l'avion (Figure 3.1). Le contrôle et l'évaluation des données issues des calculateurs primaires et secondaires est réalisée à l'aide d'un troisième type de calculateur appelé FCDC (*Flight Control Data Concentrator*) qui est lui-même redondé.

Le calculateur numérique est constitué de cartes assemblées et rangées parallèlement dans une enveloppe métallique. Chacune de ces cartes fonctionne de manière autonome et assure une fonction qui lui est propre. Généralement, il est possible de distinguer la carte gestionnaire d'entrée sortie, la carte de calcul et de traitement de données composées essentiellement de composants numériques, la carte d'alimentation électrique et éventuellement d'autres cartes telles que les protections foudres ou les cartes de navigation (FMGEC). L'architecture du calculateur est propre à une génération d'avion et ne sera pas développée dans ce manuscrit.



Figure 3.2: Calculateur de commande de vol [AIR 10]

Nous avons étudié le profil de mission relatif à un calculateur de commande de vol employé pour une application moyen-courrier qui sera appelé par la suite **profil numéro 1**.

3.2.1.2 Profil de mission électrique et thermique par phase

Le cycle de vie d'un calculateur est composé de trois phases distinctes:

- La première phase correspond à la période où l'équipement est stocké avant d'être intégré dans le système de commande de vol de l'avion. A ce stade, les composants électroniques sont assemblés mais ne sont pas alimentés. L'équipement est stocké à la température constante de 25°C.
- Lors de la seconde phase qui est une phase de tests, les équipements sont assemblés dans le système de commande de vol de l'avion et sont alimentés. L'avion est entreposé dans un hangar soumis aux températures intérieures pouvant aller de -30 à 50°C.
- La troisième phase est une phase de mission propre au modèle d'avion et à la compagnie aérienne qui l'exploite. Pour cet exemple, le temps de fonctionnement moyen du calculateur sera de 6 930 heures par an (sur 350 jours), avec une hypothèse de 3 vols par jour. Les calculateurs sont placés en zone pressurisée de manière à ce qu'ils ne subissent pas les cycles de température induits par l'altitude. Par conséquent, la température interne des calculateurs en utilisation varie entre 10°C en phase de vol et 25°C au sol.

Il est nécessaire de distinguer la température interne du calculateur de la température de jonction du composant. La température de jonction T_J est exprimée comme la somme de la température de l'environnement T_E et la température dissipée T_D selon la relation (2.1):

$$T_I = T_E + T_D \tag{3.1}$$

La dissipation thermique propre au composant varie selon la puissance consommée P et la résistance thermique jonction-ambiante du boitier d'encapsulation θ_{IA} selon l'expression (3.2) :

$$T_D = P \cdot \theta_{JA} \tag{3.2}$$

Classiquement, le fonctionnement d'un circuit intégré digital à température ambiante et sans système de refroidissement, engendre une augmentation de la température de l'ordre de quelques degrés pour une mémoire, jusqu'à plusieurs dizaines de degrés dans le cas d'un microprocesseur. Le système de ventilation de la baie avionique permet de dissiper la chaleur issue des cartes en fonctionnement. De ce fait la température de jonction des composants peut être contrôlée en régulant la température interne du calculateur. Pour cette analyse du profil de mission, une dissipation thermique équivalente à 30°C a été choisie, ce qui constitue une valeur intermédiaire entre des composants mémoires et des microprocesseurs.

Il est donc maintenant nécessaire de déterminer le cycle réel en température que subit un composant typique embarqué dans un avion lors de son utilisation journalière. La température de jonction d'un composant assemblé sur un calculateur avionique peut être déduite des données précédemment décrites au cours d'une journée type de mission (Figure 3.3).



Figure 3.3: Profil thermique N°1 de mission d'un calculateur pour une application moyen courrier

3.2.2 Autres profils de mission

De manière analogue, deux autres profils de mission spécifiques à d'autres applications aéronautiques peuvent être définis. Ces profils ne seront pas totalement détaillés dans ce manuscrit pour des raisons de confidentialité et seront appelés par la suite **profil n°2** et **3**, respectivement. Une synthèse des conditions d'utilisations thermiques et électriques de l'équipement pour les 3 profils de mission est présentée dans les Tableau 3.1 et Tableau 3.2. Il est à noter que les températures mentionnées correspondent aux températures de jonction des composants et non à la température de l'équipement.

	Phases					
Profils thermiques	Stockage Calculateur OFF	Mission Calculateur OFF	Mission Calculateur ON			
Profil nº1	25°C	15°C	40/55°C			
Profil n°2	25°C	25°C	40°C			
Profil n°3	-40/125°C	-40/125°C	125°C			

Tableau 3.1: Profils thermiques de jonction de composant assemblés dans un calculateurs utilisés pour les applications n°1, n°2 et n°3

Profils de	Р	Durée Totale	
fonctionnement	Stockage OFF	StockageMission Calculateur ON/OFFdurée totale Application	
Profil nº1	30 jours	165 900 h / 30 ans à 40°C 42 000 h / 30 ans à 55°C	~ 30 ans
Profil n°2	10 ans	29 200 heures / 10 ans	~ 20 ans
Profil n°3	24 ans	300 secondes / 1 fois	~ 24 ans

Tableau 3.2: Profils de fonctionnement électrique pour les applications n°1, n°2 et n°3

La température de jonction d'un composant assemblé sur un équipement utilisé pour l'application 3 peut varier entre -40°C et 125°C. Néanmoins, il est difficile de connaitre avec exactitude les phases de température de la mission. C'est pourquoi, nous utiliserons pour cette mission une température de jonction du composant égale à 125°C pour les 300 secondes d'utilisation. Cette valeur constituera un pire cas.

3.2.3 Spécifications en fiabilité

3.2.3.1 Définition de la fiabilité d'un composant

La fiabilité se définit comme l'aptitude d'un composant ou d'un équipement, à tenir ses spécifications de fonctionnement dans des conditions d'utilisation données et sur une durée spécifiée. Généralement en électronique, la probabilité de fiabilité associée, noté R(t), est exprimée en fonction du temps dans le formalisme de Weibull à deux paramètres par:

$$R(t) = 1 - F(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}$$
(3.3)

où F(t) est le pourcentage de défaillance, t le temps, η est le temps à défaillance moyen (soit une probabilité de défaillance F(t) de 63,2%) et β la pente de la distribution de la loi de Weibull. L'extraction de ce paramètre permet à l'utilisateur de caractériser la population des défaillances des composants par rapport à des parties du cycle de vie de la population totale, à savoir les défauts de jeunesse, de période de vie utile ou d'usure.

Si β < 1, les défaillances traduisent des mauvais procédés de fabrication. Dans cette représentation, R(t) décroit dans le temps pour se stabiliser une fois les défaillances de jeunesse éliminées.

Si $\beta = 1$, alors les défaillances sont aléatoires et indépendantes du temps. Le plus souvent ces défaillances traduisent :

- l'impact d'un défaut latent,
- la robustesse du composant vis-à-vis des perturbations extérieures telles que les chocs au niveau de l'assemblage ou les événements SEU et MBU au niveau de la puce.

Enfin, si β >1, alors les défaillances sont représentatives de l'usure naturelle des composants. L'usure est une grandeur statistique qui témoigne de l'endommagement progressif des composants au cours du temps. Ces dégradations ne sont généralement pas liées à des relations contraintes-déformations (robustesse) mais à des accumulations d'endommagement générées par le fonctionnement intrinsèque. Les mécanismes de défaillances étudiés dans le chapitre 2, à l'exception du TDDB FEOL pour des oxydes inférieurs à 3 nm, sont tous représentatifs de l'usure.

Toutefois, les niveaux de fiabilité spécifiés d'un composant ne sont jamais définis par les fabricants à l'aide de la probabilité de défaillance, car R(t) est dépendante de la durée de mission, qui est elle non connue par le fabricant. Les niveaux de fiabilité du composant électronique sont de ce fait décrits de manière indissociable:

- Le taux de défaillance aléatoire $\lambda_{\text{COMPOSANT}}$, calculé pour les mécanismes aléatoires pendant la durée de la mission (β =1). $\lambda_{\text{COMPOSANT}}$ est une grandeur probabiliste indépendante du temps, qui est exprimée en défaillance par heure-composant ou FIT (avec 1 FIT= 1 défaillance x 10⁹ / [Nombre de dispositifs x Nombre d'heures de test]). Ce paramètre traduit le taux de défaillance aléatoire pendant la durée de la mission.

- Le temps à défaillance d'usure $TTF_{COMPOSANT}$, déterminé pour les mécanismes de défaillance du composant, dits d'usure. $TTF_{COMPOSANT}$ est une grandeur statistique exprimée en heures, qui indique le temps à partir duquel les défauts d'usure deviennent prédominants sur les défauts aléatoires du circuit intégré. Ces défauts sont irréversibles et limitent par conséquent la durée de l'application. Expérimentalement, il existe statistiquement une distribution du temps à défaillance d'usure en fonction de la population de composant, qui peut s'expliquer par la variabilité des procédés de fabrication puce à puce, wafer à wafer et/ou lot à lot. De ce fait, nous choisirons comme référence le temps à défaillance moyen de la distribution pour lequel F(t) = 63,2 %.

De plus, pour aller plus loin, il est nécessaire de prendre en compte la notion de réparabilité en complément de celle de la fiabilité.

Une population de composants est considérée comme fiable lorsque son taux de défaillance est inférieur au taux de défaillance requis pour la mission ($\lambda_{COMPOSANT} < \lambda_{MISSION}$) et lorsque sa durée de vie est supérieure à la durée de la mission de l'équipement ($TTF_{COMPOSANT} > t_{MISSION}$). Avec les notions statistiques d'apparition de la défaillance, il est d'usage de tracer la courbe en baignoire du taux de défaillance (λ). Elle permet de visualiser schématiquement la notion fiabiliste du vieillissement à l'aide de l'évolution du λ de tout système, équipement ou composant, en fonction du temps d'utilisation.

Cette courbe se décompose ainsi en une période de rodage dit de jeunesse (λ décroissant) une période de maturité ou de vie utile (λ constant) puis la zone d'usure ou de vieillissement (λ croissant). Avec les définitions précédentes du temps de défaillance $TTF_{composant}$, nous pouvons visualiser Figure 3.4, la zone hachurée pour laquelle la population des composants est considérée comme fiable par rapport à $\lambda_{MISSION}$ ou à t_{MISSION}.



Figure 3.4: Schématique de la courbe du taux de défaillance d'une population de composant $\lambda_{\text{COMPOSANT}}$ en fonction du temps de fonctionnement faisant apparaitre la région de fiabilité délimitée par le taux de défaillance requis (λ_{MISSION}) pendant la durée de la mission (t_{MISSION}). L'influence du mode de vieillissement par usure (pointillés) se traduit par le décalage de la courbe en λ faisant apparaitre la croissance des défaillances plus tôt réduisant ainsi la durée de vie utile du composant.

Ce raisonnement est un cas général mais si le composant est intégré dans un équipement non réparable, le schéma précédant est simplifié. En effet, seul le critère de la durée de l'application reste pertinent quelle que soit la nature des mécanismes (aléatoires ou usure).



Figure 3.5: Utilisation dans un équipement non réparable : Positionnement de la défaillance du composant en fonction du temps, faisant apparaitre la région de fiabilité d'un composant par rapport à la durée de l'application (t_{APPLI}) et le temps moyen à la défaillance

3.2.3.2 Spécifications de fiabilité pour les 3 applications

La définition des paramètres de fiabilité nous impose de connaître la réparabilité de l'équipement pour chaque application. Dans les trois exemples d'application étudiés, seul l'équipement utilisé pour le profil n°1 en mission est réparable. Les profils 2 et 3 sont par défaut non réparables. Pour le profil n°1, le taux de défaillance requis est de l'ordre de 100 FITs soit 10⁻⁷ défaillance⁻¹.composant⁻¹. Enfin, la durée de vie requise pour un composant t_{MISSION} est donnée par la durée de mission de l'équipement. Les paramètres de fiabilité relatifs aux 3 profils sont donnés dans le Tableau 3.3.

	Réparabilité		Phases	
Spécifications en fiabilité par application	de	Stockage	Mission	
	l'equipement	t _{mission} λ _{mission}		t _{MISSION}
Application nº1	OUI	30 jours	100 FITs	30 ans
Application n°2	NON	10 ans	0 défaillance	10 ans
Application n°3	NON	24 ans	0 défaillance	300 secondes

Tableau 3.3: Spécifications en fiabilité pour les applications n°1, n°2 et n°3

3.3 Application des modèles de dégradation pour trois applications aéronautiques

Cette section se propose de calculer la durée de vie de structures élémentaires BEOL et FEOL soumises aux trois profils de missions préalablement décrits dans la section 3.2. Pour cela, les modélisations des mécanismes présentés dans le chapitre 2 au niveau silicium vont être appliquées.

(a) Pour les mécanismes de claquage franc d'oxyde (grille ou intermétallique), le temps à défaillance du composant en application TTF_{APPLI} est défini comme étant le produit du temps à défaillance obtenu dans les conditions de test TTF_{TEST} et du facteur d'accélération AF. TTF_{TEST} est obtenu à partir des résultats de tests issus de la bibliographie alors que le facteur AF est lui calculé à partir des données technologiques du composant et du profil de mission considéré.

$$TTF_{APPLI} = TTF_{TEST} \cdot AF \tag{3.4}$$

(b) Pour les mécanismes porteurs chauds et l'instabilité de la tension de seuil en température (NBTI) qui traduisent une dérive de paramètres, l'approche est quelque peu différente. Tout d'abord, les paramètres des modèles de dégradation ont été ajustés sur des points expérimentaux de la littérature et cela pour plusieurs natures d'oxydes. Ensuite, la dérive notée $\Delta P/P_o$, est calculée pour la durée de la mission. Si la dérive calculée atteint le critère de défaillance fixé, alors le temps à défaillance dans les conditions de l'application TTF_{APPLI} , est extrait. Si ce n'est pas le cas, la dérive est extrapolée jusqu'à ce qu'elle atteigne le critère de défaillance. Il est alors obtenu une durée de vie et une marge de vie pour l'application considérée (Figure 3.6).



Figure 3.6 : Méthode de calcul de la durée de vie d'un transistor MOS pour les mécanismes de dérive à partir de caractéristiques $\Delta P/P_o(t)$.

3.3.1 Les mécanismes de défaillance BEOL

Le chapitre 1 a montré que pour un nœud technologique donné, l'architecture des interconnexions peut beaucoup différer selon les fabricants de semiconducteurs. Toutefois, pour mener à bien une analyse de criticité des mécanismes BEOL, il est nécessaire de fixer certains critères technologiques. Ces critères concernent principalement la nature des matériaux, la géométrie des métallisations et des barrières de diffusion ainsi que la nature de l'oxyde d'isolation. Pour effectuer une analyse exhaustive, nous nous sommes fondés sur les données ITRS fournies entre 1998 et 2009 **[ITR 07]** pour répertorier les différents paramètres géométriques résumés dans les Tableau 3.4 et Tableau 3.5 et qui seront utilisés tout au long de la section 3.3.1.

	Para	Paramètres électriques et géométriques des métallisations en aluminium					
Noeud (nm)	Polarisation composant I/O (V)	Largeur de piste W (nm)	Epaisseur de piste h (nm)	Longueur de piste L (nm)	Epaisseur des couches de passivation δ _{CP} (nm)	Epaisseur des barrières de diffusion δ _{BD} (nm)	Permittivité de l'oxyde SiO ₂
500	5	500	1000		60	30	
350	5	350	500		50	25	
250	3.3	250	350		40	20	
180	3.3	180	250	1 1010	30	14	2.0
130	2.5	130	180	1.1010	20	10	5.9
90	2.5	90	130		15	7.3	
65	1.8	65	90		10	5.2	
45	1.5	45	65		7	3.3	

Tableau 3.4 : Paramètres géométriques des métallisations en aluminium [ITR 07]

	Pa	Paramètres électriques et géométriques des métallisations en cuivre					
Noeud (nm)	Polarisation composant I/O (V)	Largeur de piste W (nm)	Epaisseur de piste h (nm)	Longueur de piste L (nm)	Epaisseur des couches de passivation δ _{CP} (nm)	Epaisseur des barrières de diffusion δ _{BD} (nm)	Permittivité de l'oxyde Low K
130	2.5	130	180		20	10	3.5
90	2.5	90	130	1 1010	15	7.3	3.3
65	1.8	65	90	1.1010	10	5.2	3.1
45	1.5	45	65		7	3.3	2.9

Tableau 3.5 : Paramètres géométriques des métallisations en cuivre [ITR 07]

3.3.1.1 L'électromigration

Dans un circuit intégré submicronique, les métallisations globales, intermédiaires et locales sont conçues avec des dimensions différentes. De ce fait, plusieurs générations technologiques de métallisations peuvent coexister au sein d'un même empilement BEOL. Pour estimer le risque d'un composant vis-à-vis de l'électromigration, il est donc nécessaire de considérer les métallisations présentant la dérive de grain la plus importante. Les sections suivantes présentent un exemple de calcul analytique de dérive de grain, pour des composants aux métallisations en aluminium et en cuivre, soumis aux profils de mission 1, 2 et 3 et cela pour différentes générations technologiques.

3.3.1.1.1 Critères d'utilisation

Les conditions d'utilisations thermiques et électriques de fonctionnement ont été détaillées au préalable par les profils décrits dans les Tableau 3.1 et Tableau 3.2.

Pour le mécanisme d'électromigration, il est admis qu'une dérive de matière n'est possible que si et seulement si le composant est sous alimentation. Autrement dit, seules les phases de mission de l'équipement seront prises en compte pour le calcul de dérive et pas les phases de stockage. En toute rigueur, la période pendant lequel le composant est sous tension est inférieure ou égale à la période où l'équipement est en phase de mission. Pour cela, un facteur α a été introduit pour modéliser le pourcentage de la durée de sollicitation des métallisations lorsque l'équipement est en fonctionnent. Toutefois, puisque l'objectif est de connaitre la durée de vie des métallisations dans une utilisation correspondant à un pire cas, ce facteur sera fixé à 100 % dans les calculs.

Enfin, dans le cas où le profil de mission thermique est composé de plusieurs paliers de température, l'expression de la dérive de matière d_{EM} est donnée par :

$$d_{EM} = \frac{\alpha}{100} \cdot \sum_{i}^{N} d_i(T_i)$$
(3.5)

avec :

$$d_i(T_i) = v_i(T_i) \cdot t_i(T_i) \tag{3.6}$$

où $d_i(T_i)$ est le déplacement de matière calculé à la température T_i , N est le nombre de paliers, $v_i(T_i)$ la vitesse de dérive à la température T_i . L'expression de ce dernier paramètre est donnée dans le chapitre 2 par l'équation (2.10) pour les métallisations en aluminium et par l'équation (2.11) pour les métallisations en cuivre. Enfin, $t_i(T_i)$ est la durée du palier à la température T_i fournie dans le Tableau 3.2 pour les trois profils de mission.

Dans cette approche nous devons nous fixer un seuil de criticité. Les interconnexions seront considérées défaillantes lorsque le déplacement de matière d_{EM} atteint 1 µm dans le sens longitudinal. A l'aide de ce critère de défaillance, le temps à défaillance « pire cas » ainsi que la marge de vie du composant vis-à-vis de l'électromigration ont pu être extraits. Toutefois, il faut souligner qu'en dépit de la formation de cavité dans le volume de la ligne, il est possible que la

piste soit toujours opérationnelle, notamment grâce aux barrières de diffusion qui font office de conducteur électrique. C'est pour cette raison qu'en pratique l'électromigration se mesure à l'aide d'une variation de résistance relative, et non pas par une dérive de grain.

Néanmoins, l'approche proposée présente certains avantages. Le premier avantage est que les modèles de dégradation utilisés sont fondés sur la physique de la défaillance et de ce fait, ils ne nécessitent pas d'ajustements à l'aide de coefficients théoriques. Le second avantage est sans doute la possibilité de prendre en compte le nœud optimum de la technologie et la topologie des lignes et des barrières de diffusion réelles, ce qui est dans cette étude la finalité recherchée.

3.3.1.1.2 Paramètres technologiques

Les paramètres technologiques BEOL nécessaires à la détermination de l'impact de l'électromigration dans les nœuds $L_G = 0,5 \mu m$ jusqu'au 45 nm sont récapitulés dans les Tableau 3.4 et Tableau 3.5.

Les paramètres des modèles de dégradation retenus sont ceux des chemins de diffusions principaux. C'est à dire, à la frontière des grains pour l'aluminium et aux interfaces pour le cuivre. Les autres chemins sont considérés comme n'ayant peu ou pas d'influence. Les paramètres de diffusion ainsi que les énergies d'activation, pour les deux matériaux sont consignées dans le Tableau 3.6.

Paramètres technologiques	Aluminium	Cuivre	Références	
Sites de diffusion	Frontières des grains	interfaces		
$Z^*{}_{GB}$	- 4	Х	- [BAO 03]	
Z^*_{CP}, Z^*_{BD}	X	- 1		
ę (μΩ.cm)	2,65	1,67	[GIG 03]	
${ m D}_{0{ m GB}}({ m cm}^2.{ m s}^{-1})$	1.71	X		
${ m D}_{0_{ m CP}},{ m D}_{0_{ m BD}}({ m cm}^2.{ m s}^{-1})$	Х	0,78		
$E_{A}\left(eV ight)$	0.7	0.8	[BIE 04]	

Tableau 3.6: Paramètres de diffusion pour l'aluminium et le cuivre

La densité de courant à travers la ligne est également un paramètre clé du modèle. Elle se calcule à partir du courant continu I et de la section S de la métallisation traversée. Dans cette étude, une valeur de courant égale à 1 μ A (transistor CMOS HP **[ITR 07]**) a été choisie. La

section S est elle estimée comme étant le produit entre la hauteur et la largeur de la ligne. De ce fait, la densité de courant peut être déterminée pour toutes les générations technologiques.

Dans la majorité des cas, les interconnexions ne sont pas sollicitées de manière statique mais en fonctionnement digital dynamique. Comme nous ne pouvons distinguer, dans le cadre de cette étude, l'incidence de fonctionnements dynamiques analogiques, même si certaines configurations réelles font appel à des tensions de polarisations DC (montage cascode, miroirs de courants), nous nous fonderons uniquement sur un cas DC qui représentera le pire cas d'évaluation.

3.3.1.1.3 Analyse des résultats

Les résultats des calculs analytiques effectués pour déterminer les dérives ou migration d'atomes lors des contraintes en électromigration sont présentés dans les figures 3.7 à 3.12 respectivement pour les interconnexions aluminium et cuivre, pour les trois profils de mission. Le temps de défaillance correspondant est ensuite tracé dans les figures 3.13 à 3.15. L'ensemble de ces résultats permet de dresser plusieurs constats importants :

- Il est observé que les métallisations les plus étroites sont celles qui ont des durées de vie les plus courtes et cela indépendamment du choix du matériau.
- Les interconnexions en cuivre semblent être moins sensibles à l'électromigration que celles fabriquées en aluminium (Figure 3.13 et Figure 3.14). Dans le cas des applications 1 et 2, il est clairement préférable de choisir des circuits intégrés routés avec des métallisations en cuivre. La dérive de grains de cuivre n'est pas assez importante pour que celle-ci dépasse le critère de défaillance pendant la durée des applications.
- Le cas des interconnexions en aluminium est nettement plus problématique puisque en dessous des nœuds 500 nm pour l'application 1 (Figure 3.7) et de 130 nm pour l'application 2 (Figure 3.8), les spécifications ne sont plus respectées.

Les durées de vies calculées Figure 3.13 à 3.15 sont obtenues pour un pire cas, ce qui impose que la métallisation est sollicitée de manière statique lorsque le calculateur est en fonctionnement, ce qui n'est en réalité pratiquement jamais le cas. Par conséquent, la forte dérive

de grains calculée pour l'aluminium avec les figures précédentes ne peut être considérée uniquement de façon DC. Nous pouvons faire l'hypothèse qu'au cours de la mission le fonctionnement AC (digital) a un rapport cyclique $\alpha = 10$ %, ce qui se traduirait par un décalage vertical des dépendances de durée de vie Figures 3.13 à 3.15 vers le haut d'un facteur 10. Dans ce cas, cela montrerait que les dernières générations submicroniques en aluminium gravées en 90, 65 et 45 nm, ne passent toujours pas le critère de défaillance du profil 1. En revanche, un rapport cyclique $\alpha = 10$ % permettrait de fiabiliser toutes les générations de métallisation pour l'application 2.



Figure 3.7: Déplacement de grains d'aluminium en fonction du temps pour une contrainte de type électromigration, pour le profil 1, pour les nœuds 0,5 µm à 45 nm.



Figure 3.9: Déplacement de grains d'aluminium en fonction du temps pour une contrainte de type électromigration, pour le profil 2 pour les nœuds 0,5 µm à 45 nm.



Figure 3.8: Déplacement de grains de cuivre en fonction du temps pour une contrainte de type électromigration, pour le profil 1, pour les nœuds technologiques 130 nm à 45 nm.



Figure 3.10: Déplacement de grains de cuivre en fonction du temps pour une contrainte de type électromigration, pour le profil 2, pour les nœuds technologiques 130 nm à 45 nm.



Figure 3.11: Déplacement de grains d'aluminium en fonction du temps pour une contrainte de type électromigration, pour le profil 3 pour les nœuds $0,5 \ \mu\text{m} \ \text{a} \ 45 \ \text{nm}.$



Figure 3.12: Déplacement de grains de cuivre en fonction du temps pour une contrainte de type électromigration, pour le profil 3, pour les nœuds technologiques 130 nm à 45 nm.

Enfin, dans le cas de l'application aéronautique N°3, la dérive de matière est insignifiante vis-à-vis de la très courte durée d'utilisation du composant (300s), comme on l'observe Figure 3.11 pour l'aluminium et Figure 3.12 pour le cuivre. De ce fait, l'électromigration ne représente pas un risque de fiabilité majeur en termes de durée de vie (Figure 3.14) pour ce court profil de mission.



Figure 3.13: Temps à défaillance et marge de vie calculés en fonctions des nœuds technologiques 500 nm à 45 nm, pour des interconnexions en aluminium et en cuivre soumises à une contrainte de type électromigration représentative (a) du profil d'utilisation n°1, (b) du profil d'utilisation n°2.



Figure 3.14: Temps à défaillance et marge de vie calculés pour les mêmes technologies et échantillons que les figures précédentes sous électromigration correspondant au profil d'utilisation n°3.

3.3.1.2 La dérive induite par des contraintes thermomécaniques (SIV)

Dans le chapitre 2, nous avons montré que le mécanisme SIV se traduisait par deux phases, la nucléation de lacunes et la migration de grain ou d'atomes métalliques comme conséquences des contraintes thermomécaniques dans l'interconnexion. Il a été également souligné dans cette partie que la vitesse de dérive de grains pouvait s'exprimer en fonction de la température et du gradient de contrainte hydrostatique maximum de la structure avec l'équation (2.15).

Nous nous proposons dans une première partie, de déterminer par simulation la contrainte hydrostatique qui est nécessaire au calcul de la vitesse de dérive. Cette modélisation fine a été effectuée par éléments finis et a mis en évidence les contraintes mécaniques, les lignes de déformations et les chemins critiques en fonction des différents matériaux. Dans les sections suivantes, nous nous sommes servis de ces résultats pour les appliquer à la modélisation du mécanisme SIV dans les dernières générations, en calculant de manière relative et normalisée la dérive de grain lorsqu'une métallisation (*i.e.* un composant) est utilisée dans une application aéronautique.

3.3.1.2.1 Modélisation des structures aluminium et cuivre par les éléments finis

La simulation par éléments finis nous a permis de visualiser les régions sous fortes contraintes ainsi que les déformations engendrées par la différence de coefficient de dilatation thermique existant entre les matériaux qui composent la brique BEOL, lorsque la température varie. Elle rend possible une détermination précise des mécanismes impliqués dans le SIV par l'étude de la localisation des zones les plus fragiles, la quantification des contraintes ainsi que des déformations au niveau des métallisations fabriquées avec des géométries et des matériaux variés. Dans le but de distinguer plus précisément les enjeux décrits dans le chapitre 2 (section 2.3.2), nous avons effectué une analyse par éléments finis du gradient de contrainte hydrostatique.

La migration atomique sous contraintes mécaniques ou hydrostatiques impliquées au niveau BEOL, est propre au degré d'intégration d'une technologie et par conséquent, à l'utilisation des matériaux qui composent les niveaux d'interconnexions cuivre, aluminium, et alliages Al-Cu. Par conséquent, comme il n'existe pas de représentation ou de valeurs standards, il est donc nécessaire d'obtenir des estimations de l'importance des gradients de stress supposés être proportionnel aux stress mécaniques appliqués au niveau des couches minces et des effets thermomécaniques. Pour cette raison, nous nous proposons dans cette partie de calculer le gradient de contrainte hydrostatique en comparant l'utilisation d'interconnexions en aluminium et en cuivre utilisées pour des technologiques actuelles de générations 130 nm à 45 nm.

3.3.1.2.1.1 Géométrie des structures BEOL

La structure en aluminium est composée de deux lignes en aluminium (Al) connectées par un via en tungstène (W) et isolés du reste de l'empilement par un oxyde en silice (SiO₂) (Figure 3.15). Chaque ligne est recouverte (Figure 3.16) de deux barrières de diffusion en titane (Ti) et nitrure de titane (TiN) ainsi que d'une couche de recouvrement en nitrure de silicium (SiN).



Figure 3.15: Métallisation en aluminium de génération 65 nm (vue sans oxyde)



Figure 3.16: Métallisation en aluminium de génération 65 nm (vue avec oxyde)

La structure en cuivre est constituée de deux lignes de cuivre représentatives d'un procédé de fabrication double damascène. L'isolation est réalisée à l'aide d'un oxyde à basse permittivité (SiOCH) (Figure 3.17). La couche de recouvrement est également constituée de nitrure de silicium (SiN) mais les barrières de diffusion sont en tantale (Ta) et nitrure de tantale (TaN) (Figure 3.18).





Figure 3.17: Métallisation en cuivre double damascène de génération 65 nm (vue sans oxyde)

Figure 3.18: Métallisation en cuivre de génération 65nm (vue avec oxyde)

Les structures BEOL en aluminium et en cuivre ont été dessinées pour 4 géométries différentes. Ces géométries correspondent aux nœuds technologiques 130, 90, 65 et 45 nanomètres. Les paramètres principaux de ces géométries sont regroupés de manière exhaustive dans le tableau 3.7 pour les structures en aluminium, et dans le Tableau 3.8 pour les structures en cuivre.

	Paramètres géométriques des métallisations en aluminium (nm)						
Génération technologique	Largeur de piste W	Epaisseur de piste h	Longueur de piste L	Epaisseur de la couche de passivation δ_{CP}	Epaisseur des barrières de diffusion δ_{BD}	Diamètre du via	
130 nm	160	180	470	20	10	120	
90 nm	125	130	470	15	7,3	83	
65 nm	87	90	470	10	5,2	60	
45 nm	60	65	470	7	3,3	42	

Tableau 3.7 : Paramètres géométriques des métallisations en aluminium utilisés pour le calcul du gradient de contrainte hydrostatique [ITR 07]

	Paramètres géométriques des métallisations en cuivre (nm)						
Génération technologique	Largeur de piste W	Epaisseur de piste h	Longueur de piste L	Epaisseur de la couche de passivation δ_{CP}	Epaisseur des barrières de diffusion δ _{BD}	Diamètre du via	
130 nm	160	180	470	20	10	120	
90 nm	125	130	470	15	7,3	83	
65 nm	87	90	470	10	5,2	60	
45 nm	60	65	470	7	3,3	42	

Tableau 3.8 : Paramètres géométriques des métallisations en cuivre utilisés pour le calcul du gradient de contrainte hydrostatique [ITR 07]

3.3.1.2.1.2 Propriétés matériaux

Les propriétés électrothermiques et mécaniques associées à chacun des matériaux modélisés sont regroupées dans le Tableau 3.9. Certains paramètres tels que le module de Young et le coefficient de dilatation thermique peuvent varier en fonction de la température. Les valeurs consignées dans le tableau sont données pour une température de référence de 22°C.

Matériaux	Coefficient de dilatation thermique (°K-1)	Module de Young (Pa)	Coefficient de Poisson (S.D)
Al	2,30.10-5	7,10.1010	0,33
W	4,50.10-6	4,09.1011	0,28
Ti	9,40.10-6	9,60.1010	0,36
TiN	7,40.10-6	2,85.1011	0,20
SiN	8,00.10-7	3,12.1011	0,26
SiO ₂	2,61.10-6	1,30.1011	0,28
Cu	1,74.10-5	1,25.1011	0,30
SiCOH	6 , 36.10 ⁻⁵	2,50.109	0,34
Та	4,48.10-5	1,83.1011	0,35
TaN	4.70.10-5	$2.00.10^{11}$	0.30

Tableau 3.9: Principales caractéristiques des matériaux modélisés [ZHA 01], [WU 08]

3.3.1.2.1.3 Conditions aux limites et symétries

Les conditions aux limites du modèle mécanique considéré pour cette étude sont :

- Un ancrage des déplacements selon x dans le plan YZ,
- Un ancrage des déplacements selon y dans le plan XZ,
- Un ancrage des déplacements selon z dans le plan XY,

Des conditions de symétrie ont été imposées selon les plans YZ et XY.

3.3.1.2.1.4 Maillage et nombre d'éléments

Les géométries ont été maillées à l'aide d'un pavage à trois dimensions disposant chacun de 20 nœuds de calcul à interpolation quadratique. Le nombre d'éléments et de nœuds de calcul peut varier légèrement selon la technologie. Des valeurs proches de 270 000 éléments et de 1 100 000 nœuds ont dues être effectuées pour l'ensemble des géométries.

3.3.1.2.1.5 Profil thermique

Les simulations ont été réalisées pour les paliers de températures -55°C, 0°C, 25°C, 40°C, 50°C, 80°C, 100°C, 150°C, 200°C et 250°C.

3.3.1.2.1.6 Visualisation de la contrainte de Von Mises

Les résultats des simulations thermomécaniques montrent que les zones de contraintes sont principalement situées au niveau de l'interface entre la couche de recouvrement, le via et l'oxyde, et cela, quelle que soit la nature de la métallisation ou la génération technologique. Cette interface est illustrée par une flèche rouge pour les technologies aluminium (Figure 3.19c) et cuivre (Figure 3.20c). Cette observation est conforme avec les résultats obtenus de **[WU 07]** sur des structures en cuivre de génération 500 nm.

Dans les deux configurations, la contrainte est plus marquée sur le coté du via situé sur le prolongement de la ligne métallique (Figure 3.19 d et Figure 3.20 d). Ceci peut être expliqué par la faible tenue mécanique de l'oxyde d'isolation.





Figure 3.19: (a) Métallisation aluminium 65nm maillée, (b) Contrainte de *Von Mises* à 250°C sur une métallisation aluminium 65 nm vue générale, (c) vue zoomée face avant, (d) vue zoomée arrière du via





Figure 3.20: (a) Métallisation cuivre 65nm maillée, (b) Contrainte de *Von Mises* à 250°C sur une métallisation cuivre 65 nm vue générale, (c) vue zoomée face avant, (c) vue zoomée arrière du via

3.3.1.2.1.7 Visualisation de la déformation élastique équivalente

Les deux structures se déforment différemment lorsque celles-ci sont contraintes en température. Les métallisations en aluminium ont tendance à se dilater à haute température dans les 3 directions mais sont cependant bien contenues par les barrières de diffusion en titane et nitrure de titane (Figure 3.21). En revanche, les interconnexions en cuivre ne se dilatent pas mais se déplacent selon l'axe longitudinal z (Figure 3.23). Ce phénomène semble être plus prononcé pour la ligne inférieure. Ceci s'explique notamment par le fait que les deux éléments rigides de la structure (*i.e.* la couche de recouvrement et le via) se situent sur la face supérieure de la ligne alors que la face inférieure n'est maintenue que par l'oxyde à basse permittivité, qui, lui ne dispose que d'une tenue mécanique assez faible.

La déformation mécanique maximum se situe globalement aux mêmes endroits même si la structure en aluminium dispose de deux maxima. Le premier se situe à l'interface entre la couche de recouvrement, le via et l'oxyde comme pour les structures en cuivre (point 1 sur la Figure 3.22 et la Figure 3.24) alors que le deuxième maximum est plutôt localisé entre la couche de recouvrement, le via et la ligne de métal (point 2 sur la Figure 3.23). La valeur de la déformation maximale calculée pour l'aluminium est plus importante au point 2 qu'au point 1 mais reste néanmoins inférieure à celle calculée pour le cuivre au point 1. Pour les structures 65nm par exemple, la déformation calculée à contrainte équivalente est 3 fois plus importante dans les structures en cuivre que dans les métallisations en aluminium. Ce constat est également vérifié pour les nœuds technologiques 130, 90 et 45 nm.



Figure 3.21: Déformation élastique de la métallisation en aluminium 65 nm à 250°C, facteur d'échelle 30



Figure 3.22: Localisation de la déformation élastique maximum d'une métallisation en aluminium 65 nm à 250°C



Figure 3.23: Déformation élastique de la métallisation en cuivre 65 nm à 250°C, facteur d'échelle 30



Figure 3.24: Localisation de la déformation élastique maximum d'une métallisation en cuivre 65 nm à 250°C

3.3.1.2.1.8 Chemin critique

La localisation des contraintes permet de positionner un chemin de criticité. Ce chemin critique est situé selon l'axe -z à l'intersection entre le plan de section du via et du plan supérieur de la couche de recouvrement. Les contraintes normales sont considérées comme négligeables selon les axes x et y. Le chemin est composé de 200 points de calcul et est illustré par une flèche violette entre le point A et le point D (Figure 3.25). Il servira par la suite de vecteur de référence sur lequel la contrainte hydrostatique sera calculée. Les points B et C représentent les points de contact subissant une contrainte de *Von mises* maximale.



Figure 3.25: Chemin critique pour une métallisation 65 nm a) en aluminium b) en cuivre

3.3.1.2.1.9 Contrainte hydrostatique et gradient de contrainte

La visualisation de la contrainte hydrostatique sur le chemin critique montre deux comportements différents selon les deux types de technologies étudiées. Les structures en aluminium montrent deux extrema de contraintes aux points B et C (Figure 3.26), alors que les structures cuivre n'en possèdent qu'un au point B (Figure 3.27). Ces extrema sont plus prononcés pour des températures proches de 250°C. Toujours au niveau des points B et C, il existe également des extrema négatifs à -55°C mais en absolu, leurs valeurs sont inférieures à celles calculées à haute température et seront de ce fait négligées par la suite.

Pour les géométries étudiées, ces maximas sont expliqués par les singularités entre les propriétés matériaux des barrières de diffusion, de la couche de passivation ainsi que de l'oxyde d'isolation. De ce fait, il est préférable de raisonner en termes de gradient de contrainte hydrostatique plutôt qu'en contrainte seule.

La représentation du gradient hydrostatique montre cette fois ci un seul extremum pour aluminium et deux pour le cuivre comme le montre les Figure 3.28 et Figure 3.29). Toutefois, l'extremum se situe toujours aux niveaux des points d'interfaces B et C quelle que soit la technologie, les dimensions ou encore la température. En valeur absolue, pour le nœud 65 nm, la valeur du gradient maximal de contrainte est 1.6 fois plus importante dans les structures cuivre que dans les structures en aluminium. Le gradient maximum de contrainte hydrostatique extrait a aussi été tracé au niveau du point B pour les deux types de structures, en fonction du nœud technologique et de la température. Il en résulte que les dimensions et la température influent différemment selon les technologies BEOL.



Figure 3.26: Contrainte hydrostatique calculée le long du chemin critique en fonction de la température, structure aluminium 65 nm



Figure 3.27: Contrainte hydrostatique calculée le long du chemin critique en fonction de la température, structure cuivre 65 nm



Figure 3.28: Gradient de contrainte hydrostatique calculée le long du chemin critique en fonction de la température, structure aluminium 65 nm



Figure 3.29: Gradient de contrainte hydrostatique calculée le long du chemin critique en fonction de la température, structure cuivre 65 nm



Figure 3.30: Gradient de contrainte hydrostatique calculée au point B, en fonction de la génération technologique et de la température dans les structures aluminium



Figure 3.31: Gradient de contrainte hydrostatique calculée au point B, en fonction de la génération technologique et de la température, dans les structures cuivre

- Pour les technologies en aluminium, le gradient maximum de contrainte varie faiblement en fonction des dimensions. Entre les nœuds 130 nm et 45 nm, l'augmentation calculée est de l'ordre de 2, quelque soit la température. L'impact de la température n'est pas non plus très marqué pour les températures comprises entre - 55°C et 100°C. Pour les températures allant entre 100°C et 250°C, la variation de gradient maximum augmente d'un facteur 3.
- En revanche, les gradients maximums de contraintes issus des structures en cuivre, semblent être plus sensibles au nœud technologique. Pour l'ensemble des températures étudiées, une augmentation d'un facteur 4 entre les nœuds 130 nm et 45 nm a notamment été extraite. La température semble également jouer un rôle prépondérant sur le gradient de contrainte des structures en cuivre. Cette observation est surtout vérifiée pour les nœuds 45 nm où une augmentation de l'écart de gradient de 3 GPa/nm a été relevée entre -55°C et 250°C alors que pour le nœud 130 nm, cet écart représente 800 MPa/nm sur cette même gamme de température.

3.3.1.2.2 Critères d'utilisation

Le mécanisme SIV est un mécanisme de défaillance induit par des contraintes thermomécaniques. De ce fait, contrairement aux autres mécanismes de défaillance du circuit intégré, le mécanisme SIV est susceptible de se produire aussi bien dans les phases de stockage qu'au cours des phases de mission. C'est pour cette raison que les deux parties du profil de mission seront prises en considération. Les conditions d'utilisations thermiques et électriques sont données par les profils précédemment décrits dans les Tableau 3.1 et Tableau 3.2.

La dérive de grain SIV s'exprime comme la somme de la dérive en stockage d_{STOCKAGE} et de la dérive en mission d_{MISSION} , elle-même pondérée par l'inverse du facteur de sollicitation des métallisations lorsque l'équipement est sous tension ($\alpha = 100$ % dans la suite des calculs) :

$$d_{SM} = \frac{\alpha}{100} \cdot \left(d_{STOCKAGE} + d_{MISSION} \right)$$
(3.7)

Dans le cas où une phase est composée de plusieurs plateaux de températures, la dérive en stockage ou en mission, se calcule comme étant la somme des dérives singulières $d_i(T_i)$ en fonction du temps $t_i(T_i)$, où les métallisations du composant sont à la température T_i au cours de la mission. L'expression est donnée dans ce cas par la relation :

$$d_i(T_i) = \sum_i^N v_i(T_i) \cdot t_i(T_i)$$
(3.8)

où $t_{\text{STOCKAGE/MISSION}}$ est le temps total de stockage ou de mission, N est le nombre de palier et enfin v_i (Ti) la vitesse de dérive de grains à la température T_i . L'expression de ce dernier paramètre est donnée dans le chapitre 2 par l'expression (2.15).

3.3.1.2.3 Paramètres technologiques

Les paramètres technologiques BEOL relatifs aux nœuds 130 nm jusqu'au 45 nm sont disponibles dans les Tableau 3.4 et Tableau 3.5. Les valeurs du gradient de contrainte hydrostatique des métallisations, obtenues précédemment pour les technologies sont données dans la Figure 3.29 pour l'aluminium et dans la figure 3.30 pour le cuivre. Enfin, des valeurs d'énergie d'activation de 0,58 eV et 0,60 eV ont été respectivement retenues pour l'aluminium [SIV 06] et pour le cuivre [YU 07].

3.3.1.2.4 Analyse des résultats

Les dérives calculées pour les trois applications sont présentées Figure 3.32, Figure 3.33 et Figure 3.34. De façon à pouvoir être comparés entre les diverses applications, les résultats ont été normalisés par rapport à la plus petite dérive calculée, à savoir la dérive de grain d'aluminium en technologie 130 nm pour le profil numéro 2. De ces résultats, nous pouvons souligner plusieurs points importants :

Tout d'abord, les métallisations en cuivre possèdent une dérive de grain bien plus importante que les métallisations en aluminium. Ce constat est vérifié quelle que soit l'application considérée. Pour le profil numéro 2 notamment, la dérive de grain de cuivre d'une métallisation en 45 nm peut même atteindre un facteur 1,5. Ce constat s'explique par la forte différence de coefficient d'extension thermique entre l'oxyde d'isolation à basse permittivité et le cuivre (Figure 3.33).

Nous confirmons que la dérive est plus marquée pour des nœuds technologiques les plus intégrés. Cette observation est également valide quel que soit le profil de mission. Pour les trois applications, l'augmentation de la dérive entre les nœuds 130 et 45 nm est de l'ordre de 2 pour l'aluminium et d'un facteur 3,3 pour le cuivre. Ces augmentations se justifient par le fait que le gradient de contrainte hydrostatique à l'interface entre le via, l'oxyde et les barrières de diffusion, augmente avec l'intégration des technologies.

Enfin, le profil d'utilisation le plus critique semble être le numéro 3. Ceci s'explique notamment par la durée de stockage assez conséquente de l'application 3 par rapport au fonctionnement. Le profil de mission numéro 2 est plus conservateur.

Toutefois, il n'a pas encore été possible de relier ces dérives à un temps de défaillance du fait du manque de données expérimentales disponibles. Ces points sont nécessaires pour aller plus loin dans l'étude de ce mécanisme.



Figure 3.32: Dérives de grain de métaux calculées pour un composant soumis au mécanisme de stress migration pour le profil de mission numéro 1 dans les dernières technologies submicroniques.



Figure 3.33: Dérives de grain de métaux calculées pour un composant soumis au mécanisme de stress migration pour le profil de mission numéro 2 pour les technologies $L_G = 130$ nm à 45 nm.



Figure 3.34: Tracés identiques aux deux précédentes figures, appliqués au profil aéronautique numéro 3.

3.3.1.3 Le claquage d'oxyde intermétallique

La fonctionnalité d'un circuit intégré impose des espacements intermétalliques différents selon la localisation des niveaux globaux, intermédiaires et locaux. C'est-à-dire que plusieurs espacements peuvent coexister au sein d'un même empilement BEOL d'un circuit intégré. De ce fait, une bonne évaluation de l'importance de ce mécanisme impose de ne considérer que l'espacement minimal existant entre deux pistes du niveau 1. Cette considération sera prise par la suite comme postulat.

La section suivante se propose de calculer le temps de claquage de l'oxyde intermétallique pour des composants fortement submicroniques de génération 0,5 µm jusqu'au 45 nm, pour les trois applications aéronautiques considérées.

3.3.1.3.1 Critères d'utilisations

Les critères d'utilisation thermiques et électriques ont été donnés pour les 3 applications par le Tableau 3.1 et le Tableau 3.2. Le claquage de l'oxyde intermétallique n'est envisageable que si le composant est alimenté. Autrement dit, seules les phases de mission seront prises en compte dans les calculs en fonction des différentes phases thermiques 40°C/55°C (profil 1), 40°C (profil 2) puis à 125°C (Profil 3). Dans le cas où le profil de mission thermique serait composé de plusieurs paliers comme lors des phases de vols et d'escales, l'expression du facteur d'accélération du mécanisme est donnée par la relation (3.9):

$$AF_{TDDB_BEOL} = \frac{\alpha}{100} \cdot \frac{1}{t_{ON}} \cdot \sum_{1}^{N} (AF_i \cdot t_i)$$
(3.9)

où α le pourcentage de sollicitation des métallisations lorsque l'équipement est sous tension (fixé dans cette étude à 100 % pour traduire une utilisation pire cas), t_{on} est le temps total où le calculateur est en fonctionnement, N est le nombre de paliers et AF_i est le facteur d'accélération calculé à la température T_i donné par la relation (2.19).

3.3.1.3.2 Paramètres technologiques

Les paramètres technologiques BEOL relatifs aux nœuds 500 nm jusqu'au 45 nm ont été récapitulés précédemment dans les Tableau 3.4 et Tableau 3.5. Les paramètres du modèle de dégradation du mécanisme sont regroupés dans le Tableau 3.10 avec les références de la littérature. Ces paramètres vont nous servir à calculer la durée de vie des oxydes intermétalliques.

	Nature de l'oxyde			
Paramètres	SiO_{2} (k = 3.9)	SiOCH (k = 2,7)		
γ (cm/MV)	3,45	4,31 [KIM 07]		
E _A (eV)	0,7	0,3 [CHE 08]		
β	4	1,5 [HAA 07]		
TDDB _{TEST} (s)	1,105	1,10 ² [KIM 07]		
E_{TEST} (MV/cm)	5,5	3 [KIM 07]		
T_{TEST} (°C)	150	105 [KIM 07]		
S _{TEST} (cm ²)	4,1 10-4	4,1.10 ⁻⁴ [KIM 07]		

Tableau 3.10: Paramètres utilisés pour le modèle de claquage d'oxyde intermétalliques pour le SiO₂ et pour le SiOCH poreux à basse permittivité.

3.3.1.3.3 Analyse des résultats

Les résultats des calculs analytiques concernant le mécanisme de claquage de l'oxyde intermétallique montrent d'une part que les oxydes à faible permittivité sont 100 fois plus sensibles que leur équivalent en silice et d'autre part que la durée de vie diminue fortement. Toutefois, comme le montrent les Figure 3.35a,b et Figure 3.36 dans le cas des applications aéronautiques 1-3, le choix du matériau d'isolation n'a pas beaucoup d'importance puisque celuici n'affecte pas la marge de vie du composant de façon significative. Seul le profil N°3 montre un écart réduit en durée de vie qui tend à souligner le faible impact de la nature du diélectrique par rapport aux stress courts.

De plus, les résultats sont représentatifs d'un pire cas pour lequel un champ électrique est appliqué à travers l'oxyde de manière permanente. En réalité lorsque le composant est utilisé, le champ électrique n'est présent que lorsque les interconnexions avoisinantes sont traversées par un courant électrique. De ce fait, il serait pertinent de considérer pour les calculs du temps de défaillance de ce mécanisme, la probabilité qu'une interconnexion soit sollicitée électriquement par rapport à un taux d'activité d'un bloc ou dispositif fonctionnel (facteur de sollicitation \neq
100%). Ceci se traduirait par des valeurs de temps à la défaillance bien supérieures à celles calculées.

En considérant les oxydes d'isolation en silices ainsi que ceux en SiCOH (de permittivité égale à 2,7), pouvons donc observer que le claquage d'oxyde intermétallique ne représente pas un risque de fiabilité majeur pour les critères de fiabilité des applications aéronautiques citées (nœud 500 à 45 nm). Il est a noté que ce constat n'est valide que pour les oxydes étudiés. Il ne saurait être extrapolé pour différentes natures d'oxydes à basse permittivité tels que le SIOF et le SilK ou encore les oxydes SiCOH utra-poreux (k = 2,4).



Figure 3.35: Durée de vie d'un composant soumis à un claquage d'oxyde intermétallique pour la silice et les oxydes à faible permittivité (a) Pour les générations $L_G = 500$ nm à 45 nm du profil N°1, (b) pour les générations $L_G = 500$ nm jusqu'au 32 nm du profil N°2.



Figure 3.36: Durée de vie d'un composant soumis à un claquage d'oxyde intermétallique pour le profil n°3, pour la silice et les oxydes à faible permittivité des générations $L_G = 500 \text{ nm}$ jusqu'au 32 nm.

3.3.2 Les mécanismes de défaillance FEOL

De manière analogue à la section 3.3.1, les mécanismes de défaillance FEOL peuvent être analysés de manière à estimer la durée de vie des transistors du circuit intégré lorsqu'ils sont soumis aux profils aéronautiques 1, 2 et 3. Cette section permet d'évaluer la criticité des mécanismes de claquage d'oxyde de grille, de porteurs chauds ainsi que d'instabilité de la tension de seuil en température pour les différentes générations technologiques de composant.

Pour ce faire, nous avons dû fixer certains critères FEOL correspondant aux différents transistors. Les paramètres physiques et électriques correspondants aux nœuds technologiques 500nm jusqu'au 45nm, sont issus de l'ITRS **[ITR 07]** et sont regroupés dans le Tableau 3.11 en fonction de la tension d'alimentation (V_{DD}), l'épaisseur équivalente d'oxyde de grille (EOT), de la tension de seuil (V_{THo}) pour une technologie haute performance (section 1.3.2). La longueur effective est ici approximée avec une approche de type SPICE 3 où $L_{EFF} = L_G - 2 \times 0.75 \times L_D$ avec L_D la longueur de diffusion latérale des zones d'extension source et drain assimilée à la profondeur de jonction ($L_D = X_p$).

Nœud (nm)	Paramètres technologiques NMOS / PMOS W _G = 1µm							
	V _{DD} (V)	EOT (nm)	V _{TH0} (V)	L _{EFF} (nm)	L _D (nm)	X _I (nm)		
500	5	12	0,57	350	100	100		
350	3,3	9	0,45	251	66	66		
250	2,5	5	0,43	181	46	46		
180	1,6	2	0,30	129	34	34		
130	1,3	1,6	0,28	91	26	26		
90	1,2	1,3	0,27	66	16	16		
65	1,1	1,1	0,26	47	12	12		
45	1,1	1	0,25	30	10	10		

Tableau 3.11: Paramètres technologiques FEOL en fonction des différents nœuds [ITR 07]

Avec ces données géométriques et de polarisation, l'influence des différents modes de défaillance tels que le TDDB, la dégradation porteurs chauds et le NBTI sera modélisée aux tensions nominales de fonctionnement (V_{DD}) sous contrainte DC, c'est-à-dire à:

- $V_{GS} = V_{DD}$: TDDB, NBTI
- $V_{DS} = V_{DD}$: HCI au pire cas de fonctionnement, *i.e.* pour $V_{GS} = V_{DD}/2$ pour $L_G > 130$ nm et $V_{GS} = V_{DD}$ pour $L_G \le 130$ nm.

Et cela pour les 3 profils de mission, comme précédemment, c'est-à-dire en fonction des cycles OFF-ON et en cycles thermiques (Figure 3.3) correspondant aux 3 applications.

3.3.2.1 Le claquage de l'oxyde de grille

3.3.2.1.1 Application aux composants submicroniques

Le chapitre 1 a souligné la grande diversité de technologies existantes pour un même nœud technologique de composant (Tableau 1.1). Ceci est particulièrement vrai en deçà du nœud 90nm où 3 types d'oxyde de grille de transistor peuvent coexister dans les circuits intégrés du marché. La section suivante présente une simulation analytique du temps de claquage de l'oxyde de grille pour les générations CMOS 500nm jusqu'au 45nm, pour les différents types d'oxydes et pour les trois profils de mission aéronautiques. Il est à préciser que les résultats ci-dessous représentent un pire cas DC comme critère d'évaluation standard au cours duquel la fonction programmée par l'équipement sollicite toujours le même transistor qui vieillira par conséquent de manière statique et permanente. Le ratio au vieillissement dynamique AC-DC n'est pas traité ici devant la grande diversité des formes de signaux qui doit s'appliquer pour une charge donnée, ce qui nous rendrait tributaire du type de la cellule considérée, de ses interconnexions et des cycles de charge/décharge pour un nœud technologique.

3.3.2.1.1.1 Critères d'utilisation

Les conditions d'utilisations thermiques et électriques ont été préalablement spécifiées pour chaque profil de mission 1, 2 et 3 dans les sections 3.2.1 et 3.2.2. La première condition d'apparition du TDDB est bien évidemment que le claquage de l'oxyde n'est possible que lorsque le composant est alimenté. Autrement dit, seules les phases de mission sont prises en compte dans les calculs. Dans le cas où le profil de mission thermique est composé de plusieurs paliers, l'expression du facteur d'accélération du mécanisme est donnée par :

$$AF_{TDDB_FEOL} = \frac{\alpha}{100} \cdot \frac{1}{t_{ON}} \cdot \sum_{1}^{N} (AF_i \cdot t_i)$$
(3.10)

où α est le pourcentage de sollicitation des transistors lorsque l'équipement est sous tension (fixé à 100 % dans la suite des calculs pour être représentatif d'un pire cas d'utilisation), t_{ON} est le temps total où le calculateur est en fonctionnement, N est le nombre de paliers et AF_i est le facteur d'accélération calculé à la température T_i. Ce paramètre important est donné par l'équation 2.20.

3.3.2.1.1.2 Paramètres technologiques

Il n'existe pas de modèle de défaillance universel pour le claquage d'oxyde de grille dans les technologies fortement submicroniques. La distinction fondamentale sur le TDDB se situe tout d'abord au niveau du champ dans l'oxyde. De plus la large gamme d'épaisseur d'oxyde n'implique pas les mêmes mécanismes d'injections entre les oxydes épais (Fowler-Nordheim) et les oxydes tunnels ($T_{ox} \leq 3nm$) ce qui rend totalement caduque l'utilisation d'un seul modèle pour toutes les épaisseurs T_{ox}. Classiquement, le modèle thermochimique dit en E (en exponentiel du champ) est plus adapté pour les champs électriques faibles alors que le modèle dit en 1/E (exponentiel de 1/E) est plus représentatif des champs électriques forts, associé au mécanisme AHI (section 2.4.1.1.3). Cependant, il est communément admis que le modèle en E représente un pire cas par rapport à celui en 1/E [VOL 04], ce dernier étant moins conservatif. Par conséquent, le modèle en champ sera utilisé par la suite pour les simulations du mécanisme en TDDB. Il faut souligner que le modèle en puissance de la tension de grille ($t_{BD} \propto V_G^{-N}$) est généralement employé pour les oxydes tunnels (équation 2.22) montrant des valeurs pour l'exposant N \cong 44 à 40 [WU 02], [WU 05], [SUN 05], ce qui donne des résultats encore plus optimistes que les modèles exponentiels avec la réduction de V_G. Ce modèle ne sera donc pas utilisé dans ce manuscrit du fait également du manque de données disponibles dans la littérature.

Les paramètres technologiques et les conditions d'utilisations dont nous avons besoin ont été décrits par le Tableau 3.11. Les données de la littérature qui vont nous servir à effectuer nos calculs prévisionnels sont appliquées à l'aide du modèle en E (équation (2.22)) et rassemblées dans le Tableau 3.12. L'aire des échantillons est définie comme $S_{APPLI} = W_G \ge U_G$ et le champ électrique appliqué $E_{APPLI} \cong V_{DD}/EOT$ tandis que les cycles thermiques sont définis par les 3 profils aéronautiques.

NT 1		Paramètres du modèle			Données expérimentales				
(nm)	Oxyde	E _A (eV) [KIM 99]	γ (cm/MV)	β	E _{TEST} (MV/cm)	T _{TEST} (°C)	S _{TEST} (cm ²)	TDDB _{TEST} 63% (s)	Ref.
500	SiO_2	0,86	3	9	14,78	140°C	10-4	100	[WU 05]
350	SiO_2	0,82	2,96	4,5	14,78	140°C	10-4	100	[WU 05]
250	SiO ₂	0,75	2,8	3	14,78	140°C	10-4	100	[WU 05]
180	SiO ₂	0,60	2,54	1,6	14,78	140°C	10-4	100	[WU 05]
130	SiO ₂	0,45	2,24	1,3	14,78	140°C	10-4	100	[WU 05]
150	SiON	0,45	2,24	1,5/2,4	18,70	125°C	10-5	700	[CHE 01]
00	SiO ₂	0,35	1,95	1,2	14,78	140°C	10-4	100	[WU 05]
90	SiON	0,35	1,95	1,5/2,4	18,70	125°C	10-5	700	[CHE 01]
65	SiO ₂	0,31	1,76	1,2	14,78	140°C	10-4	100	[WU 05]
05	SiON	0,31	1,76	1,5/2,4	18,70	125°C	10-5	700	[CHE 01]
	SiO ₂	0,24	1,6	1,2	14,78	140°C	10-4	100	[WU 05]
	SiON	0,24	1,6	1,5/2,4	18,70	125°C	10-5	700	[CHE 01]
45	HfSiON	0,80	5,7	1,4	16,40	140°C	10-7	1000	[MCP 02]
	HfSiO	0,80	5,7	1,4	17,18	125°C	10-8	200	[MCP 02]
	HfO ₂	0,80	1,6	2	17,85	125°C	1,5.10-5	20 000	[MCP 02]

Tableau 3.12: Paramètres et données expérimentales du modèle en E utilisés pour la modélisation du TDDB en fonction des générations technologiques.

3.3.2.1.1.3 Analyse des résultats

L'analyse du TDDB en fonction de l'intégration des différentes technologies nous permet de souligner plusieurs points de la littérature. En premier lieu, la durée de vie des oxydes de grille à 63% d'apparition du claquage diminue fortement avec l'intégration et cela quelle que soit la nature de l'oxyde, *i.e.* que ce soit pour la silice ou les diélectriques Hi-K. Pour les composants fortement submicroniques composés d'oxyde de grille SiO₂, cette diminution est estimée d'un facteur 100 sur les trois dernières générations et plus largement d'un facteur proche de 10⁹ pour les générations 500nm jusqu'au 45nm. Ce constat est également vérifié pour les oxydes nitrurés où la diminution de durée de vie donne un facteur proche de 10⁵ entre les nœuds 130 nm et 45nm. Cette diminution s'explique principalement par la réduction non homothétique de l'oxyde de grille T_{ox} par rapport à la tension d'alimentation V_{DD}.



Figure 3.37: Durée de vie d'un composant soumis au claquage de l'oxyde de grille pour le profil n°1, pour la silice, les oxydes nitrurés et à haute permittivité des générations $L_G = 500$ nm jusqu'au 45 nm.

D'autre part, il est possible de comparer la sensibilité des différents oxydes de grille vis-àvis du mécanisme de claquage. Les oxydes à forte permittivité tels que le HfSIO, le HfO₂ et le HfSION présentent Figure 3.37 à Figure 3.39, une durée de vie nettement supérieure à leur équivalent en silice pour le nœud 45 nm en raison de l'effet de la constante diélectrique, notamment sur la tension de claquage. De même, nous observons que les oxydes nitrurés sont préférables depuis le nœud 130 nm. Ils permettent d'augmenter la durée de vie au claquage du composant d'un facteur 400 par rapport aux oxydes SiO₂ et cela indépendamment du profil de mission.

De manière générale, l'application des critères de profils de missions aéronautiques 1 à 3, montre que le choix de l'oxyde de grille maintenu en SiO₂ (Figure 3.37, Figure 3.38 et Figure 3.39) présente de sérieuses limitations vis-à-vis du TDDB avec l'intégration. En effet, seuls les oxydes en silice semblent être contre-indiqués qu'à partir du nœud 32 nm pour l'application n°1. C'est une des raisons pour laquelle ces oxydes ont été remplacés par la plupart des fabricants de semiconducteurs, dans un premier temps dès les générations 130/90 nm par l'oxyde SiON, puis pour les dernières générations 32nm, par le diélectrique Hi-K à base d'hafnium.



Figure 3.38: Mêmes types de tracés pour le claquage d'oxyde de grille appliqué au profil n°2.



Figure 3.39: Mêmes tracés de durée de vie TDDB des figures précédentes appliqués au profil n°3.

Etant donné que les trois profils montrent une marge de sécurité, le mécanisme de claquage de l'oxyde de grille ne représente donc pas un risque majeur pour les composants utilisés dans les applications aéronautiques des différents profils considérés.

3.3.2.2 Les porteurs chauds

Le temps à la défaillance d'un composant soumis à une contrainte de type porteurs chauds (CHC) a été calculé analytiquement pour des transistors NMOS de génération $0.5\mu m$ à 45nm constituées d'oxydes de grille SiO₂ et SiON. Ce choix est dicté par le fait que ces transistors subissent une dégradation CHC plus sévère dans les technologies CMOS sub-microniques.

3.3.2.2.1 Critères d'utilisation

Les conditions d'utilisation thermiques et électriques sont données par les profils précédemment décrits pour les missions 1 à 3. Par définition, la dégradation porteurs chauds n'est effective que lorsque le dispositif est sous tension DC (ou AC). La mise à l'état OFF ($V_{GS} = 0$), ou en latence du transfert d'un bit lors d'un transitoire sur l'entrée $V_{IN} = V_{GS}$ ($V_{DS} = 0$ à la sortie). n'implique que la relaxation vis-à-vis des phases de stress, en dehors des considérations à haute température coté PMOS où sévit le NBTI qui sera traité dans la section 3.3.2.3. Ainsi, avec les mêmes arguments développés plus haut pour ce critère, seules les phases de fonctionnement seront prises en compte en alternance DC pour le calcul de la dérive de la tension de seuil. De plus, la dérive de la tension de seuil ne représente pas un pire cas de dégradation dans les technologies avancées en comparaison des critères en courant (I_{DS}), en gain (g_m , G_D) en régime linéaire et saturé. Elle offre néanmoins une base de comparaison simple en fonction des facteurs d'intégration géométriques des technologies en lien direct avec les facteurs d'accélération en polarisation et champ électrique.

Dans notre procédure, le composant est considéré comme défaillant lorsque la dérive de la tension de seuil atteint un seuil de 20 % pour les technologies supérieures à 130 nm et 10 % pour celles inférieures ou égales à 130 nm. Ce choix arbitraire sur l'amplitude du critère est volontaire et nous permet de regagner en sévérité du critère tout en compensant partiellement la plus faible sensibilité du paramètre ΔV_{TH} par rapport aux critères en courant. En toute rigueur, rien ne garantit que 57 mV de décalage sur une tension de seuil V_{THo} = 0,57 V (L_G = 0,5 µm) n'entraine obligatoirement une défaillance de la fonctionnalité de la porte (ou de la cellule) plus tardivement qu'un décalage de 25 mV sur un V_{THo} = 0,25 V du nœud 45 nm. Cependant, ΔV_{TH} entraîne un ralentissement effectif de la réponse du transistor (cellule) qui est observé par la réduction de la fréquence de commutation. Par conséquent, II est difficile de connaitre avec exactitude le niveau de dérive critique nécessaire pour provoquer une défaillance fonctionnelle du circuit suivant les marges de bruit des états logiques 0/1, notamment par rapport à (V_{IN} + ΔV_{TH})/V_{DD}. Toutefois, ces critères garantissent un niveau de sécurité couramment employé par les fabricants de semiconducteurs et seront donc utilisés dans la suite de cette étude.

De cette manière, la dérive de tension de seuil au cours de la mission $\Delta V_{TH_{HCI}}(t)$ peut ainsi être extraite à chaque fois que le transistor est utilisé à une température T_i lors des différentes missions. Pour cela, $\Delta V_{TH_{HCI}}(t)$ est exprimée comme la somme des dérives singulières

 $\Delta V_{TH_{Ti}}(t)$, calculées pour une durée d'utilisation à la température T_i considérée avec l'expression (3.11), le tout pondérée par le facteur de sollicitation de la cellule CMOS lorsque l'équipement est en fonctionnement (α est fixé à 100 % pour être représentatif d'un pire cas d'utilisation) :

$$\Delta V_{TH_{HCI}}(t) = \frac{\alpha}{100} \cdot \sum_{1}^{N} \Delta V_{TH_{T_i}}(t)$$
(3.11)

Dans un second temps, le temps à défaillance TTF_{HCI} est extrait lorsque la dérive de la tension de seuil atteint le critère de défaillance fixé à 20 % pour les nœuds supérieurs à 130 nm (3.12) et à 10 % pour les générations DSM inférieures ou égales à 130 nm (3.13).

$$TTF_{HCI} = f [\Delta V_{TH_HCI}(t) = 20 \%]^{\frac{1}{n}} \text{ si } L_G > 130 \text{ nm}$$
 (3.12)

$$TTF_{HCI} = f[\Delta V_{TH_HCI}(t) = 10 \%]^{\frac{1}{n}} \text{ si } L_G \le 130 \text{ nm}$$
 (3.13)

où f est la fonction de la dérive temporelle de la tension de seuil du transistor NMOS pour une technologie donnée et n la pente temporelle de dégradation du mécanisme HCI.

3.3.2.2.2 Paramètres technologiques

Le critère d'évaluation en mode porteurs chauds (CHC) est historiquement lié aux performances d'intégration des technologies numériques depuis les années 80. Généralement ce mode de défaillance est lié à l'usure des dispositifs au même titre que l'électromigration et le BTI. Les procédures de durée de vie sont initialement fondées sur l'accélération de la dégradation CHC sous contraintes DC de manière à gagner en temps de test, en choisissant les conditions de polarisations représentatives des pires cas de génération de défaults (section 2.4.2.1.2). Cependant, devant la sévérité de l'accélération en tension (en champ latéral) de ce mode de défaillance pour les technologies submicroniques, il est nécessaire d'appliquer un facteur de transfert DC-AC (ou *derating*) pour passer des conditions de contraintes DC aux conditions AC réelles de fonctionnement. Ceci est d'autant plus justifié que les extractions de TTF établies par rapport au critère de durée de vie DC de 10 ans sont très loin d'être satisfaisantes suivant les types de technologies et qu'il est donc nécessaire de gagner en marge (réaliste) de sécurité en multipliant par le facteur d'accélération DC-AC.

L'approximation quasi statique considère que si la forme des signaux est conservée, à travers le ratio du temps montée r sur la période du signal T_{AC} , noté $r = t_r / T_{AC}$, ainsi que la durée de la période CHC (T_{HOT}) au cours du transitoire, c'est-à-dire où $x_N = t_r / T_{Hot}$, alors le facteur d'accélération s'obtient comme :

$$AF = \frac{T_{AC}}{T_{HOT}} = \frac{T_{AC}}{t_r} \cdot \frac{t_r}{T_{HOT}} = \frac{x_N}{f_{AC} \cdot t_r} \iff NTF = \frac{x_N}{f_{AC} \cdot t_r} = \frac{x_N}{f_{AC} \cdot r \cdot T_{AC}} = \frac{x_N}{r}$$
(3.14)

où $x_N = 4$ dans le NMOS ($X_P = 10$ dans le PMOS) est le facteur employé Figure 3.40a par rapport aux technologies $L_G > 130$ nm (pire cas de dégradation dans le transitoire $V_{GS}=V_{DS}/2$) **[QUA 94]**. Ainsi, avec cette approche traditionnelle, on peut constater que le facteur AF = NTFpour les transistors NMOS (PMOS) est indépendant de la fréquence si la forme du signal r est conservée entre les nœuds technologiques. Ceci peut donner par exemple un facteur de forme du temps de montée compris entre r = 1/20 (5 %) et 1/10 (10 %), conduisant à un facteur AC-DC en temps NTF= (x_N/r) = 80 à 40 ce qui est relativement faible.



Figure 3.40 (a) Forme du signal en fonctionnement digital pour la détermination des coefficients DC-AC au cours de la dégradation porteurs chauds. (b) Dépendance du facteur d'accélération NTF en fonction du temps de montée t_r à fréquence fixée $f_{AC} = 25$ MHz, puis en fonction de la fréquence pour $t_r = 0,5$ ns fixé **[QUA 94]**.

Dans la réalité, cette approche simplifiée néglige l'influence du délai entre l'entrée et la sortie ainsi que la plus forte dépendance de t_r / T_{AC} vis-à-vis de la charge en sortie, en fonction des géométries des transistors PMOS (charge) et NMOS (décharge) où généralement $W_P = 2,5$ à 3 W_N . Par conséquent, pour une fréquence donnée, la variation de t_r et donc du facteur r, montre Figure 3.39b que le NTF effectif varie entre un facteur 20 à 340 pour une fréquence fixée à 25 MHz ($L_G = 0,5 \mu m$). Si l'on fixe $t_r = 0,5$ ns, l'augmentation de la fréquence montre que le facteur NTF varie d'un facteur 640 à 100. En pratique, le calcul s'effectue en

fonction de la fréquence par rapport à la forme du signal vue par le transistor (Figure 3.40a) donnant par exemple un NTF proche de 120 à une fréquence de 100 MHz pour un temps de montée $t_r = 0,35$ ns, *i.e.* pour r = 3.5 % [QUA 94]. Dans un exemple type pour le fonctionnement digital où r = 8 %, nous avons un facteur DC-AC, NTF = 50. Généralement, la mesure expérimentale d'un critère en DC au niveau d'un transistor est comparée aux stress AC sur le même transistor (et même polarisation) et dans des oscillateurs en anneau en fonction des ratios géométriques des étages de l'oscillateur. Ceci permet d'ajuster le facteur de forme et obtenir un ratio DC-AC réaliste plus précis, qui permet de calculer en AC la durée de vie au plus proche des conditions réelles de fonctionnement.

L'autre effet à considérer dans les technologies récentes ($L_G < 130$ nm) est que le pire cas DC n'est plus au milieu du transitoire en $V_{in} = V_{GS}$ (Figure 3.40a) mais décalé vers les hautes tensions $V_{IN} \ge V_{OUT}$ ce qui réduit la durée de la période T_{Hot} augmentant fortement le facteur x_N avec les relations 3.16. Par conséquent, même si la dégradation dans ces technologies persiste malgré la réduction de V_{DD} , les facteurs DC-AC s'en trouvent fortement augmentés [**BRA 05**], [**GUE 08**].

Les valeurs des ratios DC-AC des travaux récents de Guérin [GUE 08] ont montré que dans les technologies HP (LP) 45 nm (Figure 4.41), ces coefficients décroissent légèrement avec l'augmentation de V_{DD} . En associant ces différents V_{DD} utilisés pour les nœuds CMOS antérieurs au 45 nm, nous pouvons observer que le facteur DC-AC augmente fortement avec la réduction de la longueur depuis le 130 nm.

De plus, nous avons Tableau 3.13, des valeurs nettement plus importantes que les ratios DC-AC des technologies à canal plus long ($L_G > 130$ nm) fonctionnant à plus hautes tensions V_{DD} comme discuté précédemment. Ceci peut être attribué au fait que pour les technologies $T_{OX} < 2$ nm (effet tunnel), l'injection de porteurs s'effectue directement à partir du canal de conduction (CHC) et de ce fait, une partie des pièges générés devient relaxable. Ces données ont été obtenues à température ambiante (25°C) et seront supposées constantes pour les températures des trois applications étudiées, pour les nœuds CMOS $L_G \leq 130$ nm.



Figure 3.41: Dépendance des facteurs DC-AC pour la technologie 45 nm en fonction de la tension d'accélération V_{DD} [GUE 08].

	Coefficients DC-AC							
Nœuds (nm)	500	350	250	180	130	90	65	45
Facteurs AC-DC	250	275	300	350	3570	3645	3720	3720

Tableau 3.13: Coefficients DC-AC pour le mécanisme HCI tels que pour $L_G > 130$ nm, le ratio est calculé pour un facteur de forme r = 2.5 % (FANOUT = 1) et pour $L_G \le 130$ nm avec les données de la Figure 3.41 aux V_{DD} correspondants **[GUE 08]**.

3.3.2.2.3 Analyse des résultats

L'application du modèle PTM-HCI présenté à la section 2.4.2.2.4 permet de calculer simplement la dérive de V_{TH} avec l'équation (2.35). Les résultats des figures 3.42 à 3.47 montrent que la diminution des dimensions comme la longueur de grille L_G (W_G) et l'épaisseur du diélectrique (T_{ox}) impacte fortement le décalage de la tension de seuil du transistor NMOS qui augmente au cours du temps. Ceci conduit à une réduction importante des performances en courant qui se répercute à travers le délai ($\propto C_L V_{DD}/I_{On}$) comme un ralentissement de la fréquence de commutation de la porte en logique, puis de la fréquence globale de la réponse en fréquence d'une fonctionnalité ou du bloc considéré.

Comme nous l'avons souligné au chapitre 2, la dégradation de la structure avec le temps est liée à la génération des états d'interface suivant un exposant n ~ 0,46 (0,5) qui, selon le modèle PTM-HCI, est considéré invariant pour une même nature d'oxyde et pour chaque technologie étudiée. Ceci implique qu'il n'y a pas d'effet de saturation des cinétiques à long terme, et par conséquent à profil de mission équivalent au cours du mode ON. La pente de la dégradation des paramètres transistors reste identique dans les transistors NMOS 0,5 µm et 45 nm. Ceci traduit une signature caractéristique de ce mode de dégradation en fonction du temps de contrainte, qui a été vérifié expérimentalement à basse et haute fréquences de fonctionnement **[GUE 08]**, **[BRA 09]**.



Figure 3.42: Décalage de la tension de seuil à V_{DD} en fonction du temps de fonctionnement au cours d'un stress HCI pour le profil 1, dans les différents nœuds technologiques à oxyde de grille en SiO₂



Figure 3.44: Même critère de dégradation HCI pour le profil 2, à oxyde de grille en SiO₂



Figure 3.46 : Mêmes tracés de durées de vie HCI pour le profil 3, dans les oxydes de grille en SiO₂.



Figure 3.43: Dérives de la tension de seuil suite à une contrainte HCI à V_{DD} pour le profil 1, en fonction du temps de fonctionnement et pour différents nœuds technologiques à oxydes de grille en SiON



Figure 3.45: Même critère de dégradation HCI pour le profil 2, à oxydes de grille en SiON



Figure 3.47: Mêmes tracés de durées de vie HCI pour le profil 3, dans les oxydes de grille en SiON.

En considérant des critères de défaillance fixés respectivement à $\Delta V_T = 20$ % et 10 % pour les nœuds de 0,5 µm à 180 nm d'une part et de 130 nm à 45 nm d'autre part, on observe pour les oxydes SiO₂ plusieurs comportements. Tout d'abord, les technologies strictement supérieures à 130 nm, ne tiennent pas les spécifications en durée de vie du profil 1. Pour ces nœuds, la durée de vie sous injections de porteurs chauds se situe dans une gamme de 21-27 ans, soit en dessous de la spécification avionique requise qui est de 30 ans (Figure 3.48).

Ensuite, un gain de durée de vie est observé pour les nœuds sub-130 nm. Ceci s'explique tout d'abord par la réduction de V_{DD} et la plus faible sensibilité de V_{TH} pour ces technologies visà-vis des autres paramètres transistor. De plus, c'est également dû à l'augmentation des facteurs DC/AC qui est 10 fois plus importante pour ces technologies (Tableau 3.13) et se traduit par une nette augmentation de la fiabilité AC. Néanmoins, la durée de vie AC (Figure 3.48) de ces 4 technologies récentes diminue successivement d'un facteur 2 entre chaque nœud. Ceci est observé au point de ne plus respecter les spécifications avionique du profil 1 où la durée de vie est estimée à 21 ans pour le 45 nm.

A l'inverse, il apparait que le mécanisme porteurs chauds (Figure 3.49a et Figure 3.49b) fondé sur ΔV_T n'impacte pas la mission des composants électroniques pour les applications 2 et 3, indépendamment de la nature de l'oxyde de grille SiO₂ ou SiON, qui imposent une durée d'utilisation de 10 ans et 5 minutes, respectivement.



Figure 3.48: Durée de vie en fonctionnement (AC) d'un transistor NMOS sous contraintes HCI du profil n°1 compte tenu des facteurs de formes AC/DC pour les mêmes critères (ΔV_{TH}) que les Figures 3.41-3.46 dans les générations technologiques 0,5 µm à 45 nm, pour l'oxyde de grille SiO₂ et nitruré SiON.

Cependant, il faut souligner que les oxydes thermiques en SiO₂ pur utilisés pour la grille ont pratiquement disparu de la fabrication des circuits intégrés dès les nœuds 130-90 nm. Leurs remplaçants nitrurés NO, NO₂ (thermique ou plasma) présentent une meilleure résistance aux injections de porteurs chauds, spécifiquement vis-à-vis de la génération des défauts d'interface. On observe en effet dans les Figure 3.49a,b que le gain en durée de fonctionnement (AC) au cours des profils 2 et 3 est de l'ordre d'un facteur 90 pour le nœud 45 nm.



Figure 3.49: Mêmes tracés que la figure 3.47 pour le critère de durée de vie $\Delta V_{TH} = 20 \%$ si L_G > 130 nm et $\Delta V_{TH} = 10 \%$ si L_G ≤ 130 nm (a) pour le profil n°2, (b) et pour le profil N°3.

3.3.2.3 L'instabilité de la tension de seuil en température (BTI)

La spécificité historique de ce mode de dégradation est qu'il est beaucoup plus important dans les transistors PMOS polarisés sous tension de grille négative à haute température (NBTI), notamment à travers l'amplitude de la dérive de la tension de seuil, en comparaison à la dégradation HCI où ΔV_{TH} est beaucoup moins sensible, dû à la localisation des défauts au drain du transistor. Par conséquent, le cas des transistors NMOS sous contraintes PBTI ne sera pas traité.

Le temps à la défaillance d'un transistor PMOS soumis à une contrainte de type NBTI a été calculé analytiquement pour les 3 profils de mission précédents dans les nœuds technologiques $0,5 \ \mu m$ à 45 nm pour les oxydes SiO₂.

3.3.2.3.1.1 Critères d'utilisation

Les conditions d'utilisation thermiques sont prépondérantes pour le mode de dégradation NBTI et tout particulièrement les phases sous tension (mode ON) par rapport au mode OFF. Puisque nous recherchons les effets sur la durée de vie dans les conditions pire–cas de fonctionnement, les 3 profils de missions sont étudiés en mode ON. De plus, il a été considéré pour ces simulations que le transistor PMOS est vieilli de façon DC lorsque l'équipement est sous tension (α =100 %). Dans le cas d'un point mémoire SRAM, ceci signifie que le bit écrit et stocké, reste inchangé au cours de la mission (PMOS de charge toujours ON). En pratique cela n'est jamais le cas, ce qui implique qu'en absence de contrainte électrique (état OFF où -V_{GS} < -V_{TH}) la tension de seuil du PMOS peut se relaxer. De ce fait, le cas étudié constitue réellement un pire cas.

Les énergies d'activations en températures utilisées sont comprises entre $E_A = 0,13$ eV et 0,18 eV (oxyde SiO₂ et SiON) et 0,11 eV pour les diélectriques Hi-K. Avec la même démarche que celle utilisée pour le mécanisme de porteurs chauds, la dérive de la tension de seuil d'un transistor PMOS soumis à une contrainte NBTI est donnée par l'expression (3.15):

$$\Delta V_{TH_NBTI}(t) = \sum_{1}^{N} \Delta V_{TH_Ti}(t)$$
(3.15)

De même, le temps à défaillance TTF_{NBTI} est extrait lorsque la dérive de la tension de seuil atteint le critère de défaillance fixé à 20 % pour les nœuds supérieurs à 130 nm (3.16) et à 10 % pour les générations inférieures ou égales à 130 nm (3.17).

$$TTF_{NBTI} = g[\Delta V_{TH_NBTI}(t) = 20\%]^{\frac{1}{n}}$$
 si L_G > 130 nm (3.16)

$$TTF_{NBTI} = g \left[\Delta V_{TH_NBTI}(t) = 10 \% \right]^{\frac{1}{n}} \text{ si } L_{G} \le 130 \text{ nm}$$
 (3.17)

où g est la fonction de la dérive temporelle de la tension de seuil du transistor PMOS pour une technologie donnée et n la pente temporelle de dégradation du mécanisme NBTI.

3.3.2.3.1.2 Paramètres technologiques

Les paramètres du modèle PTM BTI multi cycle que nous avons utilisés pour ces simulations sont consignés dans le Tableau 3.11. Ces paramètres ont été déterminés pour l'utilisation DC durant laquelle la polarisation de grille est imposée de manière continue dans les transistors PMOS. Cependant, dans le cas d'application en fonctionnement digital réel où le transistor vieillit uniquement dans l'intervalle de temps où il est adressé, i.e. lors de la charge Figure 3.40a où V_{IN} passe de V_{DD} à 0 (V_{GSP} = V_{IN} - V_{DD} passant à l'état bas), nous pouvons considérer que dans le cas d'une commutation de porte logique standard le transistor est 50 % du temps dans le mode NBTI, soit plus précisément, au cours de la phase temporelle de proportion (T_{AC}/2 – t_r – t_f) / T_{AC}. C'est pourquoi, nous devons considérer pour les extrapolations la dépendance en rapports cycliques AC pour la dégradation NBTI. Le pire cas statique est obtenu avec une valeur de $\alpha = T_{DC}/T_{AC}$ égal à 99,99 % (Chapitre 2, Figure 2.30) et le cas le plus court, représentatif d'un régime impulsionnel, avec $\alpha = 0,01$ %, alors que le cas typique NBTI en alternance est pour 50 %.

3.3.2.3.1.3 Analyse des résultats

La dégradation NBTI est connue comme pour être fortement dépendante de l'épaisseur de l'oxyde de grille et de la tension d'alimentation [HUA 05]. Trois natures d'oxydes de grille ont été étudiées sous contrainte NBTI :

- les oxydes en silice (SiO₂) entre les générations 500 et 45 nm,
- les oxydes nitrurés (SiON) pour les nœuds 130 à 45 nm
- les oxydes à haute permittivité de type HfSiON où seul le nœud 45 nm sera étudié.

Les exemples que nous avons étudiés sont représentatifs de technologies de transistors PMOS issus de filières hautes performances (faibles V_{TH} , fort I_{ON}). De ce fait, ces résultats ne sont extrapolables pour les autres filières basse consommation (LP) à structures équivalentes, qu'en modifiant les principales valeurs de V_{DD} , V_{TH} ; EOT(T_{ox}) et L_{eff} dans le Tableau 3.11 ainsi que les valeurs de champs électriques correspondant.

Dans le cas des oxydes SiO_2 , il est observé à l'aide des Figures 3.50, 3.52 et 3.54 que la durée de vie des transistors PMOS entre les nœuds 500 nm et 45 nm, diminue nettement avec la réduction de L_G. Cependant, cette réduction n'est pas homogène et certains nœuds n'obéissent pas à cette règle. C'est notamment le cas des nœuds 0,35 µm et 0,25 µm pour lesquels la durée de

vie est supérieure à celle des transistors PMOS issus du nœud 0,5 µm. Ce gain de fiabilité peut s'expliquer par la forte diminution de la tension d'alimentation V_{DD} qui a eu lieu pour ces deux filières (respectivement de 3,3 V et 2,5 V) et qui a conduit à des champs électriques $E_{eff} \cong V_{DD}/EOT$ plus faibles à travers l'oxyde de grille.

Ainsi, d'une manière générale, on note que quel que soit le nœud étudié, les oxydes en SiO_2 ne représentent pas un risque en fiabilité majeur pour les trois applications aéronautiques. Toutefois, seuls les nœuds 45 et 65 nm sous stress DC, ne semblent pas passer les spécifications en durée de vie de l'application 1 (Figure 3.50).

Le cas des oxydes nitrurés SiON est en revanche plus problématique (Figure 3.51, Figure 3.53 et Figure 3.55). L'étude de la criticité du mécanisme NBTI pour ces oxydes doit se traiter profil par profil. Dans le cas de l'application 1, les composants de génération 45 nm ne tiennent pas la spécification en durée de fonctionnement (Figure 3.51), si ces composants contiennent un transistor PMOS qui est adressé plus de 1,1 % du temps dans le calculateur au cours de la mission, soit plus de 2 290 heures. Ceci est obtenu par extrapolation de la dépendance de la TTF en fonction du facteur β pour les différentes longueurs et le profil de mission N°1 qui suivent une loi linéaire en TTF $\propto 1/\beta$. Pour les générations 65 nm, 90 nm et 130 nm (Figure 3.51), le pourcentage critique est extrapolé respectivement à 5,6 %, 11 % et 57 % d'alternances, soit respectivement pour 11 640 h, 22 870 h et 118 500 h de fonctionnement sur 30 ans.

Pour le profil 2, tous les nœuds technologiques tiennent les spécifications en durée de vie (Figure 3.53) même si le transistor PMOS de la puce est adressé à 99,99 % du temps sur les 10 ans d'utilisation.

Etant donné la durée très courte de l'application 3, la dérive de la tension de seuil du transistor PMOS avec un oxyde SiON (Figure 3.55) reste très faible. De ce fait, le mécanisme NBTI n'est pas problématique pour cette application en intervention ultra courte.



Figure 3.50: Durée de vie des transistors PMOS pour le profil n°1 soumis à une contrainte NBTI en fonction des nœuds technologiques 500 nm jusqu'au 45 nm HP et pour plusieurs fréquences de sollicitation pour la silice SiO₂.



Figure 3.52: Durée de vie sous contrainte NBTI d'un transistor PMOS en fonction des nœuds 500 nm à 45 nm HP pour le profil n°2, pour la silice SiO2, en fonction de la fréquence.



Figure 3.51: Durée de vie des transistors PMOS pour le profil n°1 soumis à une contrainte NBTI pour les nœuds 130 nm à 45 nm HP et pour plusieurs fréquences de sollicitation, pour les oxydes nitrurés SiON



Figure 3.53: Durée de vie sous contrainte NBTI d'un transistor PMOS en fonction des nœuds 130 nm à 45 nm HP pour le profil n°2, pour le diélectrique en SiON, en fonction de la fréquence.



Figure 3.54: Mêmes tracés de durée de vie dans les transistors PMOS 500 nm à 45 nm HP sous stress NBTI pour le profil n°3 pour la silice SiO₂.

Figure 3.55: Mêmes tracés de durée de vie dans les transistors PMOS 130 nm à 45 nm HP sous stress NBTI pour le profil n°3 pour le diélectrique SiON.

Le cas des oxydes de grille à forte permittivité de type HfSiON montre Figure 3.56 une tenue en durée de vie 10 fois moins bonne pour le nœud 45 nm que les oxydes en silice. Le profil de mission n°1 n'est pas garanti pour une durée d'utilisation de 30 ans si les transistors PMOS sont sollicités à plus de 1,3 % soit 2 700 h (Figure 3.56). Aussi, le passage au diélectrique Hi-K ne montre aucune amélioration de la durée de fonctionnement par rapport aux oxydes nitrurés puisque pour ces oxydes, les pourcentages d'utilisation critiques sont du même ordre de grandeur (environ 1.1% sur la Figure 3.50).



Figure 3.56: Durée de fonctionnement d'un oxyde 45 nm HfSiON HP soumis aux profils de mission 1, 2 et 3 pour différentes fréquences d'utilisation lorsque le calculateur est en fonctionnement.

Pour les applications 2 et 3, les oxydes HfSiON ne sont pas contre indiqués puisque même sous contrainte DC (99,99 % d'utilisation), leurs durées de fonctionnement respectives (Figure 3.56) sont supérieures aux spécifications de fonctionnement de 30 ans (Profil 1), 10 ans (Profil 2) ou 5 mn (Profil 3).

3.3.3 Comparaison des critères de défaillances

Les examens des différents modes de défaillance et de l'importance des mécanismes impliqués prépondérants au niveau du circuit intégré ont été effectués dans les sections précédentes de ce chapitre. En considérant les trois applications aéronautiques ainsi que le critère de défaillance respectif pour chaque mécanisme, un calcul de la durée de fonctionnement globale d'un composant électronique sensible aux mécanismes de défaillances du silicium peut être obtenu et resitué par rapport aux spécifications d'une application aéronautique réelle.

Le premier constat issu de cette analyse, est que la durée de vie des semi-conducteurs diminue lorsque le nœud technologique est réduit, et cela quel que soit le choix des matériaux utilisés. La pente de la diminution n'est pas uniforme pour tous les mécanismes. La diminution doit être repositionnée par rapport à une spécification de durée de vie d'application réelle pour évaluer si le mécanisme est critique ou non. Dans le cas de l'application N°3 (Figure 3.57), le TDDB induit une diminution de la durée de fonctionnement de l'ordre de 10 décades. Cependant, compte tenu de la durée de fonctionnement très courte de cette application, l'aggravation des mécanismes avec la diminution du pas technologique n'a aucune incidence. Il est en revanche plus critique pour les applications 1 et 2 ci-après.



Figure 3.57: Estimation de la durée de fonctionnement d'un composant semi conducteur CMOS soumis au profil de mission n°3 en fonction du nœud technologique, pour les mécanismes de défaillance BEOL et FEOL. Métallisations aluminium, oxyde intermétallique en SiO2, oxyde de grille en SiO2, substrat massique.

Le deuxième constat est que le choix d'un matériau ou d'une solution technologique modifie considérablement la pente de la diminution de la durée de vie. De ce fait, un mécanisme peut devenir plus actif si un matériau est remplacé dans la structure et l'ordre de criticité des mécanismes qui en découle, peut s'en trouver modifié. Ceci est notamment le cas pour l'application N°2 (Figure 3.58) où l'électromigration demeure le mécanisme le plus critique pour des structures fabriquées avec des interconnexions en aluminium, des oxydes intermétalliques et les oxydes de grille en silice. En revanche, pour les composants constitués de métallisations Cu avec des isolations en diélectrique low-k et des oxydes de grille SiON, c'est le mécanisme NBTI (Figure 3.59) qui demeure le plus limitant car le plus sévère.



Figure 3.58: Tracés de durée de fonctionnement d'un composant semi conducteur CMOS pour les mécanismes de défaillance DC BEOL et FEOL du profil de mission n°2 en fonction du nœud technologique : Métallisations Al, oxyde intermétallique en SiO₂, oxyde de grille en SiO₂, substrat massique.



Figure 3.59: Tracés de durée de fonctionnement d'un composant semi conducteur CMOS soumis au profil n°2 en fonction du nœud technologique, pour les mécanismes de défaillance DC BEOL et FEOL. Métallisations cuivre, oxyde basse permittivité, oxyde de grille en SiON, substrat massique.

Le troisième point est que la modification des conditions d'utilisation, à savoir la température de l'environnement et la fréquence d'utilisation, peut prolonger la durée de fonctionnement du circuit intégré selon le mode de défaillance considéré. Pour cela, il est d'usage de résonner à partir du temps à défaillance du circuit intégré entier plutôt que sur les temps à défaillance issus des mécanismes pris indépendamment. Le temps à défaillance du circuit intégré entier peut être stimulé avec (3.18) en sommant les durées de fonctionnement individuelles des 5 mécanismes considérés comme indépendants.

$$TTF_{IC} = \left(\frac{1}{TTF_{EM}} + \frac{1}{TTF_{TDDB_BEOL}} + \frac{1}{TTF_{TDDB_FEOL}} + \frac{1}{TTF_{HCI}} + \frac{1}{TTF_{NBTI}}\right)^{-1}$$
(3.18)

Par exemple, les composants industrialisés en procédés 65 nm, avec des métallisations en aluminium, des oxydes d'isolation interligne ainsi que des oxydes de grille en SiO₂, ne satisfont pas la spécification en fonctionnement du profil 1 s'ils sont utilisés de manière statique pendant la durée de la mission, soit entre 55°C en phase d'escale et 40° en vol (flèche rouge dans la Figure 3.60). Cependant, en abaissant le pourcentage d'utilisation à 5.3 % dans les mêmes conditions thermiques, ou encore si la température de jonction contrôlée entre 37°C et 22°C et que le rapport d'alternance est imposé à 10 % (flèche rouge sur la Figure 3.61), alors le composant peut tenir la spécification de la mission aéronautique comme le montre les valeurs du Tableau 3.14.

Il est donc possible d'établir un gabarit de bon fonctionnement du composant en fonction du pourcentage d'utilisation (Figure 3.60) et d'une gamme de température (Figure 3.61). Les plages relatives de fonctionnement du profil 1 pour les générations 65 et 90 nm sont données respectivement dans les Tableau 3.14 et 3.15. Les résultats en rouge font apparaître quand le seuil de durée de bon fonctionnement n'est pas satisfait par rapport au profil de chaque mission.



Figure 3.60: Durée de fonctionnement d'un composant soumis au profil de mission n°1 pour plusieurs pourcentages d'utilisation, en fonction du nœud technologique : Métallisations aluminium, oxyde basse permittivité, oxyde intermétallique en SiO₂, oxyde de grille en SiO₂, substrat massique



Figure 3.61: Durée de fonctionnement d'un composant soumis au profil de mission n°1 avec 10% d'utilisation, pour plusieurs gamme de température de jonction au sol/en vol, en fonction du nœud technologique : Métallisations aluminium, oxyde basse permittivité, oxyde intermétallique en SiO2, oxyde de grille en SiO2, substrat massique

Τ° / α	Pourcentage d'utilisation AC-DC a							
	99,99 %	50 %	10 %	1 %	0,01%			
55 / 40°C	1.6	3.2	16	160	16 00			
45 / 30 °C	1.9	3.8	19	190	19 000			
35 / 20 °C	3.8	7.6	38	280	280 000			
25 / 10°C	2.3	12	58	580	580 000			

Tableau 3.14: Estimation de la durée de fonctionnement d'un composant CMOS de génération 65 nm HP soumis au profil n°1, pour différents pourcentages d'utilisation AC-DC du circuit (α), en fonction des différentes gammes de température en vol et au sol. Métallisations aluminium, oxyde basse permittivité, oxyde intermétallique en SiO₂, oxyde de grille en SiO₂, substrat massique.

Τ° / α	Pourcentage d'utilisation AC-DC α							
	99,99 %	50 %	10 %	1 %	0,01%			
55 / 40°C	3.0	6.1	31	300	30 000			
45 / 30 °C	3.6	7.3	37	360	36 000			
35 / 20 °C	7.2	14	72	720	720 000			
25 / 10°C	11	22	110	1 100	1 100 000			

Tableau 3.15: Estimation de la durée de fonctionnement d'un composant CMOS de génération 90 nm HP soumis au profil n°1, pour différents pourcentages d'utilisation AC-DC du circuit (α), en fonction des différentes gammes de température en vol et au sol. Métallisations aluminium, oxyde basse permittivité, oxyde intermétallique en SiO₂, oxyde de grille en SiO₂, substrat massique.

3.4 Conclusion

Dans ce chapitre, nous avons effectué le transfert des principaux modes de défaillance détaillés dans les chapitres 1 et 2 à trois exemples d'applications aéronautiques typiques représentatives de conditions réelles de fonctionnements en vol et au sol. Les profils de mission fonctionnels et thermiques des ces 3 applications ont été détaillés à travers leurs spécifications en cyclages On/Off et en cyclages à basse et haute températures, pour laisser apparaitre les enjeux de fiabilité susceptibles de jouer un rôle important lors de ces missions.

Pour cela, une évaluation précise de la criticité BEOL et FEOL des 6 principaux mécanismes de défaillance a été réalisée avec les modèles analytiques développés au chapitre 2 et appliquée aux 3 profils de missions aéronautiques. Ceci nous a permis d'extraire les temps à défaillance médians issus de ces différents mécanismes à l'aide de données constructeurs ou connues par la littérature. La dépendance avec les paramètres géométriques principaux liés à l'intégration, ainsi qu'avec les dernières avancées technologiques utilisées pour la fabrication des récents nœuds sub-microniques décrites dans le chapitre 1, a souligné la criticité de certains modes de défaillances comme le NBTI, l'Electromigration et même les porteurs chauds (cas du diélectrique en SiO₂ pour les canaux très courts 45 nm). Dans ce dernier exemple, cela a justifié le passage au diélectrique SiON pour ces dernières générations CMOS à substrat de silicium massif, ou comme le passage aux diélectriques Hi-k pour les longueurs L_G < 40 nm. A la lumière des ces comparaisons, Il ressort plusieurs éléments qu'il est intéressant de souligner. Outre le fait que la durée de vie d'un dispositif semi-conducteur diminue avec la réduction des dimensions en longueur (L_{eff}), il apparait que l'accélération de ce phénomène traduisant une nette réduction de la fiabilité globale du composant reste très dépendante des matériaux utilisés en fonderie. Ce chapitre a permis d'identifier pour une application aéronautique contraignante, les matériaux à choisir pour respecter les spécifications en durée de vie. De ce fait, il est possible d'estimer la durée de vie d'un composant commercial DSM en identifiant les types de matériaux utilisés par les fabricants. Cette approche offre les bases de la méthodologie de sélection de composants développé en fin de chapitre 4.

De plus, ce chapitre a proposé des solutions de compromis dans le cas où un ou plusieurs temps à défaillance calculés, ne respecterait pas les spécifications des profils aéronautiques. Nous avons montré en effet qu'il est possible d'adapter certaines conditions d'utilisations telles que la température de l'environnement ou la fréquence de fonctionnement de manière à prolonger la durée de vie d'un produit.

Enfin, l'approche proposée dans ce chapitre, nous a permis d'identifier pour une application donnée, les mécanismes de défaillance les plus critiques et de ce fait, d'aiguiller l'utilisateur potentiel sur les tests à effectuer ou à examiner dans un rapport de qualification. Cette approche sera également développée dans le chapitre 4.

La méthodologie proposée repose sur de nombreux critères d'utilisation et hypothèses technologiques. Ceci signifie que les résultats de durée de vie présentés dans ce chapitre ne sont valides que pour les profils de mission et les technologies décrits en début de chapitre. D'un point de vue applicatif, si la description d'un profil change (par exemple 2 vols quotidiens au lieu de 3 dans l'application 1 ou encore un composant assemblé non plus dans un calculateur situé en baie avionique mais dans les ailes), l'estimation du temps à défaillance sera différente. D'un point de vue technologique, les dimensions BEOL et FEOL choisies pour cette étude sont issus de la filière haute performance de l'ITRS. C'est-à-dire, que les résultats présentés ici nécessitent une adaptation en dimensionnement, tension de fonctionnement, pour les autres filières existantes, comme celles à *basse consommation* et *basse consommation en veille*.

De plus, si on considère comme évoqué dans le chapitre 1, que pour un même nœud technologique, les dimensions caractéristiques du transistor peuvent différer selon les fabricants. Il devient alors difficile pour un utilisateur d'estimer la durée de vie d'un composant DSM pour son application. Par conséquent, de façon à pouvoir répondre aux besoins des utilisateurs de composants, l'approche qui consiste à sélectionner un composant par des éléments de la technologie en regard des différentes utilisations a été développée sous forme de logiciel propre à EADS.

Toutefois, cette méthodologie est fondée sur des lois de dégradation qui présentent quelques limitations. Tout d'abord, la validité des modèles sur de longues périodes est toujours discutable. En effet, ces lois sont élaborées sur des périodes très courtes de l'ordre de l'heure dans le meilleur des cas. De ce fait, il existe toujours un risque de divergence du modèle si celles-ci sont extrapolées sur de longues durées.

Ensuite, il existe également une incertitude concernant les paramètres des modèles de dégradation. Ces paramètres se traduisent le plus souvent sous forme de plage de valeur. Dans cette étude, seule la plage inférieure a été considérée de façon à être le plus représentatif possible d'une application « pire cas ». Mais, en toute rigueur, il serait préférable de traduire les courbes de dégradation présentées en intervalle de dégradation minimum-maximum.

Enfin, les lois de dégradation ne peuvent prendre en compte tous les procédés de fabrication des fondeurs. En effet, les différentes techniques de dépôt d'oxydes (plasma, ALD..), de traitement, de passivation et de recuits sont propres à un fabricant et à un composant, et impactent également l'estimation de la durée de vie.

De ce fait, afin de minimiser les incertitudes décrites plus haut, il est nécessaire de procéder à des tests de vieillissement sur des composants fonctionnels et sur de longues durées. Ceci permettra tout d'abord d'affiner la validité des modèles, et ensuite de transposer ces modèles issus du niveau silicium à un niveau produit. C'est cette démarche qui est proposée dans le chapitre 4.

4 Demonstration, validation experimentale et Methodologique pour la selection de composants

4.1 INTR	RODUCTION	173
4.2 TECH	HNIQUE EXPERIMENTALE DE SELECTION DE COMPOSANTS	173
4.2.1 V	Vieillissement HCI et NBTI sur une memoire SRAM	174
4.2.1.1	1 HCl au niveau d'un point mémoire SRAM	174
4.2.1.2	2 NBTI au niveau d'un point mémoire SRAM	175
4.2.2 D	DESCRIPTION DE LA TECHNOLOGIE DU COMPOSANT DE TEST	176
4.2.3 D	DESCRIPTION DES TECHNIQUES DE VIEILLISSEMENT	179
4.2.3.1	1 Caractérisation expérimentale d'une SRAM à l'aide de nano pointes	179
4.2.3.2	2 Test expérimental d'une SRAM à l'aide de plot FIB (Focus Ion Beam)	183
4.2.3.3	3 Banc de vieillissement de SRAM	186
4.2.4 E	BILAN	199
4.3 SELE	ECTION DE COMPOSANT PAR LES DONNEES DE TEST FABRICANT	200
4.3.1 N	Метноде	200
4.3.1.1	1 Technologie du composant	200
4.3.1.2	2 Profil de mission	201
4.3.1.3	3 Matrice de risques	201
4.3.1.4	4 Calcul du taux et du temps minimum de bon fonctionnement du composant déterminé	és dans
	les conditions de test	202
4.3.1.5	5 Facteur d'accélération (AF)	207
4.3.1.6	6 Calcul du taux et du temps minimum de bon fonctionnement du composant dans les con	ditions
		200
4.3.2 A	APPLICATION DE LA METHODOLOGIE SUR DEUX MEMOIRES FLASH NAND	210
4.3.2.1	1 Technologie des composants flash	210
4.3.2.2	2 Profil de mission	211
4.3.2.3	3 Analyse de criticité	211
4.3.2.4	4 Test de fiabilité	212
4.3.2.5	5 Calcul du temps minium de bon fonctionnement en test	213
4.3.2.6	6 Facteur d'accélérations	213
4.3.2.7	7 Calcul du temps minium de bon fonctionnement en application	214
4.3.3 S	Synthese	215
4.4 CON	ICLUSIONS	215

4.1 Introduction

Le chapitre précédent a permis de dresser une analyse de risque au niveau du silicium en fonction d'une application et d'une technologie de composant. Si cette approche est utile pour identifier les technologies à risque ainsi que les mécanismes de défaillance inhérents à une utilisation, elle nécessite toutefois d'être validée au niveau d'un composant commercial. Dans cette optique, ce chapitre vise à fournir à un utilisateur potentiel, des techniques et des outils d'analyse et d'évaluation plus précis de la fiabilité d'un composant, non plus au niveau du silicium mais au niveau du produit. Pour ce faire deux méthodes ont été identifiées.

Une première méthode consiste à déterminer le temps à défaillance du composant à l'aide de tests de vieillissements accélérés. Pour ce faire, nous présentons dans ce chapitre trois techniques de vieillissements mises en œuvre aux cours de ces travaux sur des mémoires de type SRAM. Les deux premières techniques expérimentales sont novatrices et ont été réalisées pour la première fois à des fins exploratoires dans ce domaine et pour ce type de composant. La troisième technique, plus conventionnelle à permis de mettre en pratique deux protocoles opératoires représentatifs de contraintes HCI et NBTI, spécialement développés pour cette étude.

La seconde méthode est plus orientée vers l'exploitation de résultats de tests réalisés par les fabricants au niveau de différents produits. Elle permet à un utilisateur de connaitre le temps minimum de bon fonctionnement (sans probabilité d'apparition de défaillance) d'un composant sans pour autant effectuer des tests de vieillissements. Un exemple d'application a été présenté au travers d'une mémoire flash NAND de génération 50 nm.

4.2 Technique expérimentale de sélection de composants

Le chapitre précédent a permis d'identifier les mécanismes de défaillance les plus prépondérants pour 3 applications aéronautiques identifiées. Parmi ces mécanismes, figurent l'électromigration, les porteurs chauds (HCI) ainsi que l'instabilité de la tension de seuil en température (BTI). Le risque de défaillance par électromigration peut être fortement réduit voire neutralisé, en choisissant d'effectuer le routage du circuit intégré avec des interconnexions en cuivre. De ce fait, les risques de défaillance se focaliseront ainsi sur les deux mécanismes prédominants précités de dégradation de la tension de seuil qui reste le premier paramètre pour distinguer l'état logique lors du fonctionnement.

4.2.1 Vieillissement HCI et NBTI sur une mémoire SRAM

La détermination d'un temps de défaillance fondé sur l'usure d'un composant en mission, nécessite de procéder à des tests de vieillissement accélérés (tension, température). Si les techniques de vieillissement sont nombreuses au niveau silicium, elles demeurent plus rares et plus complexes à mettre en œuvre au niveau d'un produit. En effet, il est nécessaire de prendre en compte à ce niveau le fonctionnement réel du produit, les limitations induites par le boitier d'encapsulation ainsi que par les métallisations BEOL, ainsi que les paramètres accessibles et représentatifs de ses performances. Toutefois, en appliquant les contraintes électriques et thermiques adéquates, il est alors possible d'effectuer un vieillissement accéléré de type HCI ou NBTI sur des composants SRAM commerciaux. La description précise de ces conditions opératoires est proposée dans les deux sections suivantes.

4.2.1.1 HCI au niveau d'un point mémoire SRAM

Les mémoires statiques SRAM (*Static Random Access Memory*) sont de nos jours encore largement utilisées malgré leur coût, en raison de leur rapidité et faible consommation par rapport aux mémoires dynamiques (DRAM). Elles sont principalement employées dans les applications embarquées, les mémoires caches et les tampons, pour des temps d'accès faibles ou pour la basse consommation. Ceci a pour origine le fait que lors de la mémorisation des données par des bascules, il n'y a pas besoin de rafraîchir périodiquement son contenu comme dans les DRAM. Cependant, contrairement aux DRAM, ces types de mémoires restent en permanence sous alimentation pour garder les bits stockés.

Le vieillissement HCI se caractérise sur un NMOS par l'application simultanée d'une tension grille-source V_{GS} et d'une tension drain-source V_{DS} positives. Dans une configuration CMOS, ces conditions ne sont réunies que lorsque l'inverseur est en phase de basculement lorsque la commande en entrée passe de V_{SS} à V_{DD} (Figure 4.1). Au niveau du point mémoire SRAM, le changement d'état se traduit par une opération d'écriture. La Figure 4.2 illustre un exemple d'écriture à '0' dans une cellule mémoire composée de 6 transistors, où la donnée initialement stockée en mémoire passe d'un état '1' à un état '0'. Au cours de cette transition, le transistor d'accès M6 ainsi que le transistor NMOS de décharge M3 se trouvent contraints en HCI. Si la donnée écrite change, alors ce sont les transistors M5 d'accès et M6 de décharge qui se retrouvent contraints par les porteurs chauds.

Chapitre 4 : Démonstration, validation expérimentale et méthodologique pour la sélection de composants



Figure 4.1: Conditions de vieillissement HCI sur un inverseur CMOS



Figure 4.2 : Conditions de vieillissement HCI sur une cellule SRAM à 6 transistors

De ce fait, le vieillissement d'une cellule mémoire par porteurs chauds se résume à réécrire de manière séquentielle et répétée, un état logique ainsi que son contraire par commutation continue. La dégradation se trouve même accélérée si cette séquence est opérée à basse température en raison des propriétés physiques évoquées dans le chapitre 2 (section 2.4.1.1) caractérisées par l'augmentation du phénomène d'ionisation par impact à basse température. Cette configuration de fonctionnement (tension, température) traduisant une forte aggravation de ce mode d'usure sera identifiée comme *un pire cas*, et sera retenue dans la suite de cette étude pour les vieillissements de type HCI.

Il est également possible d'accélérer le mécanisme HCI en appliquant au niveau de la cellule, une tension de polarisation supérieure à la tension nominale si l'architecture et le design de la cellule ont été prévus à cet effet (prise de contact intra cellule). Toutefois, cela reste en pratique difficilement réalisable dans les produits standards, du fait de la présence des régulateurs de tension à l'intérieur du circuit qui sont connectés aux cellules.

4.2.1.2 NBTI au niveau d'un point mémoire SRAM

La contrainte NBTI intervient dans le transistor PMOS lorsqu'une tension grille-source V_{GS} négative est imposée à travers l'oxyde de grille et quand dans le même temps, la différence de potentiel entre la source et le drain V_{DS} reste nulle. Ce type de contrainte apparait dans une structure CMOS lorsque la commande d'entrée atteint l'état logique '0' (Figure 4.3). Corollairement, l'écriture puis le stockage d'un '0 ' logique dans une mémoire SRAM 6T permet de générer une contrainte NBTI. Dans l'exemple décrit dans la Figure 4.4, c'est le transistor PMOS de charge P4 qui est contraint. Si un '1' logique est écrit puis stocké, alors le transistor P4

se relaxe et c'est le transistor P2 qui devient contraint. Afin de minimiser les effets de relaxation, la cellule ne doit pas être réécrite avec une donnée différente.

En conséquence, le vieillissement NBTI d'une cellule mémoire intervient en écrivant puis en stockant une donnée sur une longue période sans jamais en réécrire une différente. De plus, c'est l'augmentation de la température extérieure due à la dissipation thermique et la température intra dispositif qui permettent d'accélérer le mécanisme. Ce pire cas de fonctionnement sera appliqué par la suite pour vieillir les mémoires SRAM dans les conditions NBTI.

Ainsi, comme pour le mécanisme de défaillance HCI, la tension de polarisation V_{IN} de la cellule est un facteur de dégradation prépondérant pour le mécanisme NBTI.





Figure 4.3 : Conditions de vieillissement NBTI sur un inverseur CMOS

Figure 4.4 : Conditions de vieillissement NBTI sur une cellule SRAM à 6 transistors

4.2.2 Description de la technologie du composant testé

Le composant choisi pour cette étude est une mémoire SRAM de génération 90 nm, de densité 4 Mbits qui est testée dans la gamme d'utilisation industrielle (-40°C/85°C). Une analyse de construction réalisée dans les laboratoires de recherche d'EADS a permis de dresser une carte d'identité du composant au niveau BEOL et FEOL. Cette analyse est présentée dans les sections suivantes. Pour des raisons de confidentialité, la référence de ce composant ne sera pas communiquée.

La connexion électrique entre la puce et les broches est assurée par un pontage de fils en or de diamètre 20 µm. Le circuit intégré est connecté à l'environnement extérieur à l'aide de 44 broches latérales en ailes de mouette (TSOP44) (Figure 4.5). Enfin, la puce de surface 6 mm², est

maintenue sur une embase de cuivre à l'aide d'une colle argent, le tout enrobé dans une résine de type époxy d'un mm d'épaisseur.



Figure 4.5: SRAM 90 nm 4 Mbits, inspection optique

La puce est recouverte par une couche de passivation Si_3N_4 d'épaisseur 1,2 µm. L'empilement BEOL est constitué de 2 niveaux de métallisations. Les deux niveaux sont en aluminium et sont recouverts par une couche de recouvrement en titane d'épaisseur 10 nm pour le niveau M2 et de 15 nm pour le niveau M1. De même ces lignes sont supportées par un alliage titane tungstène d'épaisseur 30 nm.



Figure 4.6: SRAM 90 nm 4 Mbits, coupe micrographique AA', microscope électronique à balayage FEG, électrons secondaires, x 18 000

Figure 4.7: SRAM 90 nm 4 Mbits, coupe micrographique BB', microscope électronique à balayage FEG, électrons secondaires, x 18 000

Les vias d'interconnexions de l'empilement sont également fabriqués en tungstène, alors que les oxydes d'isolation sont réalisés en silice SiO₂. Le demi-espacement mesuré entre deux lignes du niveau M1 indique qu'il s'agit d'un circuit intégré gravé en 90 nm. Il est intéressant de faire apparaitre la spécificité des connexions de la cellule SRAM qui est constituée des 6 transistors MOS, dont deux transistors d'accès de type NMOS (N_5 et N_6 sur la Figure 4.8) car son identification à première vue n'est pas évidente. Nous avons fait également apparaitre la décomposition du point mémoire réalisé à l'aide des connexions des deux inverseurs CMOS rétro couplés. L'inverseur CMOS s'identifie par l'association classique du transistor PMOS de charge (P_3 ou P_4) et des transistors NMOS de décharge (N_1 ou N_2). La connexion électrique entre les deux inverseurs est assurée par un pontage en tungstène.



Figure 4.8: SRAM 90 nm 4 Mbits, plan mémoire 2D, microscope électronique à balayage FEG, électrons secondaires, x 40 000



Figure 4.9: SRAM 90 nm 4 Mbits, plan mémoire 3D, microscope électronique à balayage FEG, électrons secondaires, x 50 000



Figure 4.10: SRAM 90 nm 4 Mbits, plan mémoire 3D et zones de dopage p, microscope électronique à balayage FEG, électrons secondaires, x 30 000



Figure 4.11: SRAM 90 nm 4 Mbits, plan mémoire 3D et zones de diffusion n et p, microscope électronique à balayage FEG, électrons secondaires, x 40 000

Les transistors disposent de grille en poly silicium de longueur 90 nm et de largeur 150 nm. De plus, les oxydes de grille sont nitrurés (SiON) avec une épaisseur électrique équivalente de 2,3 nm. On remarque dans la Figure 4.9 les interconnexions correspondantes aux lignes de

données (BL) et de mots (WL) ainsi que les lignes d'alimentation et de masse de la cellule SRAM ainsi formée. De plus, l'observation des zones de dopages n et p, ont permis de situer les lignes d'alimentation V_{ss} et V_{DD} (Figure 4.10 et Figure 4.11).



Figure 4.12: SRAM 90 nm 4 Mbits, coupe BB' de transistors PMOS, microscope électronique à balayage FEG, électrons secondaires, x 80 000



Figure 4.13: SRAM 90 nm 4 Mbits, coupe AA' de transistors NMOS, microscope électronique à balayage FEG, électrons secondaires, x 200 000

4.2.3 Description des techniques de vieillissement

La dégradation par usure (*wearout*) est généralement longue à mettre en évidence sur des composants électroniques à leur tension nominale de fonctionnement. Aussi, de façon à minimiser les heures de tests, il est souhaitable d'appliquer conjointement tous les facteurs de contraintes en même temps, à savoir : les instructions électriques pire cas décrites par les sections 4.2.1.1 et 4.2.1.2, la température et la polarisation. Dans cette optique, la première campagne de test s'est orientée vers la caractérisation de mémoires SRAM à l'aide de nano pointes qui permet de juxtaposer les 3 contraintes.

4.2.3.1 Caractérisation expérimentale d'une SRAM à l'aide de nano pointes

4.2.3.1.1 Description

La technique nano-probing consiste à placer des nano pointes sur les contacts d'un transistor d'un composant électronique dans le but de le caractériser. Cette technique permet d'appliquer librement une tension de polarisation directement sur une structure active en évitant la limitation des régulateurs de tension internes au composant. De plus, il est possible de superposer une température de l'ordre d'une centaine de degré à l'aide d'une platine chauffante. L'appareil utilisé est un microscope à force atomique (AFM) associé à 4 pointes en tungstène de section 50 nm.



Figure 4.14: Exemple de caractérisation à l'aide de 4 nano pointes [CIM 09]

Figure 4.15: Observation optique d'une nano pointe usée, x 6

Une campagne d'évaluation de la technique nano-probing a été effectuée dans les laboratoires du Centre Intégré de Microélectronique de Provence Alpes Cotes d'Azur (CIMPACA), sur le site de Rousset. L'objectif principal était d'adapter le dispositif de caractérisation afin de pouvoir réaliser des tests de vieillissement sur des durées de l'ordre de l'heure. Dans le cas présenté, la structure de test est le transistor NMOS de décharge noté N_1 dans la Figure 4.8.

La préparation de l'échantillon est une étape essentielle à la réussite de l'expérience. Tout d'abord, la puce a été extraite du boitier par voie chimique. Ensuite les deux premiers niveaux de métaux ont été enlevés sur toute la surface par polissage mécanique au grain 50 nm. Enfin, une solution *Buffer Etch* destinée à décaper les oxydes d'isolation et à mettre en relief les contacts des cellules, a été appliquée sur la surface de la puce. De ce fait, seuls les contacts des cellules demeurent apparents.

Les 4 pointes viennent balayer indépendamment le plan des contacts afin de reconstituer une image par microscopie à force atomique. Ensuite, 3 nano pointes notées A, B et C sont positionnées respectivement sur les contacts correspondant à la grille, à la source et au drain du transistor NMOS N_1 (Figure 4.16).
Chapitre 4 : Démonstration, validation expérimentale et méthodologique pour la sélection de composants



Figure 4.16: Positionnement des nano pointes sur la grille, la source et le drain du transistor N1 de la cellule SRAM 6T, image AFM

4.2.3.1.2 Relevé des caractéristiques électriques

Deux faisceaux de caractéristiques I_{DS} - V_{GS} et I_{DS} - V_{DS} ont pu être obtenus avant vieillissement, pour le transistor N_1 issu de technologie 90nm (T_{ox} = 2,3 nm) conçu pour une tension d'alimentation V_{DD} = 1,8 V. Ces tracés sont présentés respectivement dans les Figure 4.17 et Figure 4.18.

La caractéristique I_{DS} - V_{GS} (Figure 4.17) a permis d'extraire une tension de seuil égale 0,65 V, ce qui signifie que la mémoire est issue de la filière basse consommation. La Figure 4.17 montre une dégradation de la caractéristique pour des tensions de grille supérieures 1,6 V. Cette diminution du courant n'est pas liée au composant lui-même mais plutôt à la dégradation des pointes de mesures. La détérioration des nano pointes pourrait provenir de la technique AFM utilisée initialement pour reconstituer une image. L'autre piste à envisager est celle de la densité de courant qui traverse les pointes. En effet, la densité de courant calculée pour une pointe en tungstène, de section 50 nm assimilée à un cercle, est de l'ordre de $J_{ing} \sim 9,5 \,\mu A/\pi r^2 = 4,83 \, 10^5$ A/cm^2 . Aussi, la forte densité de courant à l'extrémité des pointes a pu induire un échauffement local et donc une courbure des pointes qui perdent peu à peu leur contact.



Figure 4.17: Caractéristique IDS(VGS) du transistor N1 de la cellule SRAM 6T non vieillie

La valeur du courant de drain en saturation mesuré à l'aide de la caractéristique I_{DS} - V_{DS} est de l'ordre de 90 µA ($W_G = 2L_G = 0,18$ µm). Cette valeur est en accord avec la technologie 90 nm basse consommation. De même que précédemment, les caractéristiques du transistor en régime saturé montrent une légère dégradation pour des tensions de grille supérieures à 1,2 V (Figure 4.18), c'est-à-dire pour les plus forts niveaux de courant I_{DS} . Cet effet tend à démontrer un auto échauffement des pointes qui perdent en qualité du contact et se dégradent nettement. Du fait de cette dégradation rapide des nano pointes qui s'accélèrent d'autant plus aux conditions de stress (HCI et NBTI ces dernières nécessitent des polarisations V_{DS} (et V_{GS}) = 1,1 à 1,35 V_{DD}), il n'a pas été possible de procéder à des tests de vieillissement sous nano pointes.



Figure 4.18: Faisceau de caractéristique $I_{DS}(V_{DS})$ pour différents V_{GS} (de V_{SS} à V_{DD} = 1,8 V par pas de 0,2 V), du transistor N1 de la cellule SRAM 6T non vieillie

4.2.3.1.3 Synthèse

La technique du test à l'aide de nano pointes est utile pour déterminer les caractéristiques de fonctionnement ainsi que pour extraire les paramètres électriques des transistors tels que la tension de seuil V_{TH} et le courant de drain en saturation I_{Dsat} (I_{ON}). En revanche, il n'est pas envisageable d'utiliser ce dispositif pour des études de vieillissement du fait de la faible durée de vie des pointes (estimée à 3 minutes dans le meilleur des cas **[FOR 09]**).

Enfin, cette technique expérimentale est difficilement réalisable pour un utilisateur final du fait de son coût et de la sensibilité à la variabilité introduite par la préparation de l'échantillon et la mesure.

4.2.3.2 Test expérimental d'une SRAM à l'aide de plot FIB (Focus Ion Beam)

4.2.3.2.1 Description

Etant donné qu'il s'est avéré impossible d'utiliser des nano pointes sur une cellule SRAM de génération 90 nm pour y opérer un vieillissement, nous nous sommes orientés sur une seconde campagne de tests à l'aide d'une autre technique expérimentale : le test à partir de plots de contacts conçus par sonde ionique focalisée (FIB). L'objectif étant toujours d'appliquer conjointement les stress électriques en accélération pire cas décrits par les sections 4.2.1.1 et 4.2.1.2, la température et la polarisation. Dans le cas présenté ici, cette technique consiste à placer des plots FIB sur les contacts de la cellule SRAM sur lesquels sont posées des pointes de mesures en tungstène de taille macroscopique.

Cette campagne d'essais s'est déroulée dans les locaux du Centre National d'Etudes Spatiales (CNES) à Toulouse avec la collaboration de l'équipe d'analyse de défaillance de Thalès. Le principal objectif était d'évaluer la faisabilité de la technique pour réaliser des tests de vieillissement de type NBTI et HCI, sur une cellule SRAM 6T, en appliquant le protocole décrit dans la section 4.2.1.

Comme pour la technique de caractérisation à l'aide de nano pointes, la préparation des échantillons de test est une étape critique et difficile à réaliser. Tout d'abord, la puce a été extraite du boitier par voie chimique. Ensuite le protocole de préparation fut le suivant :

- Les deux niveaux métalliques de la mémoire ont été enlevés lors d'une étape de gravure FIB, sur une surface correspondant à 96 x 96 µm²,
- Deux tranchées ont été creusées en surface jusqu'à atteindre les contacts des cellules mémoires. Ces deux tranchées ont permis de localiser la cellule de test selon les axes x et y, en considérant la répétitivité et la symétrie du motif (Figure 4.19),
- Une fois la cellule SRAM de test localisée, les tranchées ont été recouvertes à l'aide d'une couche d'oxyde SiO₂ d'épaisseur 1 μm,
- La cellule de test a été isolée des transistors d'accès ainsi que des autres cellules SRAM à l'aide de gravures successives FIB et de dépôt d'oxyde SiO₂,
- Quatre contacts de tungstène noté A, B, C et D on été déposés respectivement sur les sources communes aux deux NMOS de décharge N₁ et N₂, sur l'entrée de l'inverseur 2 (sur la sortie de l'inverseur 1), sur l'entrée de l'inverseur 1 (sur la sortie de l'inverseur 2) et enfin sur la source commune aux deux inverseurs PMOS P₃ et P₄ (Figure 4.19),



Figure 4.19: Localisation de la cellule SRAM de test, image FIB



Figure 4.20: Réalisation de contact sur la cellule SRAM de test, image FIB



Figure 4.21: Surface de puce SRAM préparée à l'aide de plots FIB, image FIB 96 x 96 µm²



Figure 4.22: Dispositif de contrainte et de mesure électrique à l'aide de pointes macroscopiques

 Pour finir les quatre contacts préalablement déposés ont été déportés de la cellule de façon à pouvoir poser des pointes de contrainte et mesure électriques (Figure 4.20). Pour cela des plots de tungstène de surface 10 x 10 µm² ont été réalisés (Figure 4.21).

4.2.3.2.2 Résultats et bilan

Le test de SRAM à l'aide de plot FIB n'a pas permis d'obtenir des résultats cohérents du fait de la variabilité induite par la préparation des échantillons. De plus, le positionnement des contacts sur la cellule selon l'axe z, reste spatialement très délicat car rien ne garantit qu'il y ait réellement un bon contact électrique. De même, le risque de court circuit entre les contacts est omniprésent du fait de la présence de résidus post gravure.

Enfin, cette technique est très lourde à mettre en place en termes de temps de préparation et de mise en place, de matériel et de coût. De ce fait, cette technique expérimentale n'est pas conseillée pour un utilisateur final qui souhaiterait évaluer la fiabilité d'un composant

Les techniques expérimentales à l'aide de nano pointes ou de plots FIB n'ont pas permis de mettre en place des techniques de vieillissement en combinant de manière simultanée, les instructions électriques, la température et la polarisation. Par conséquent nous avons choisi de nous orienter vers une technique plus conventionnelle à l'aide d'un testeur de composant commercial que nous allons décrire ci-dessous.

4.2.3.3 Banc de vieillissement de SRAM

L'application des contraintes électriques et thermiques représentatives des vieillissements HCI et NBTI a nécessité le développement de banc de test spécifiques dédiés. La section suivante décrits les moyens de test électriques et thermiques utilisés au cours de cette troisième campagne de test.

4.2.3.3.1 Description du banc

4.2.3.3.1.1 Contrainte électrique

La SRAM est pilotée à l'aide d'un testeur de composant développé dans les laboratoires d'EADS IW à Suresnes. Le testeur ELECTRE (*Elaborated Electronic Component Tester for* $R_{eliability}^{adiation} Effects$) est composé de l'association d'une carte mère contrôlée par un FPGA avec une carte fille sur laquelle est brasé le composant de test, en l'occurrence la SRAM. Les instructions envoyées à la mémoire se font à partir de programmes de tests eux même exécutés depuis une interface graphique (Figure 4.23).

De plus, une sortie SMA (*SubMiniature version A*) permet de réaliser un suivi de la consommation électrique du composant pendant la phase de vieillissement. Le testeur est présenté dans la Figure 4.24.



Figure 4.23: Ecran de l'interface graphique ELECTRE



Figure 4.24: Testeur de composant ELECTRE

4.2.3.3.1.2 Système de contrainte thermique locale (Imhotep)

Les tests en température destinés à mettre en évidence les mécanismes de défaillance au niveau silicium, ne sont pas une chose aisée. En effet, il est impératif de s'assurer que la température n'induit pas de défaillance à un autre endroit que dans la puce silicium. Pour cela, les techniques qui consistent à placer le testeur entier dans une étuve climatique ne sont pas satisfaisantes dans le sens où un des composants de la carte mère peut être défaillant bien avant le cœur du composant à tester rendant la détection non représentative de la fiabilité du composant lui même. D'autre part, les autres configurations dans lesquelles la carte fille serait déportée dans une enceinte thermique ne conviennent pas non plus. Car en effet, plus la longueur des câbles de connexions entre la carte mère et la carte fille, est importante, plus le temps de propagation des signaux devient important. Ceci a pour effet de limiter la fréquence des instructions issues de la carte mère et induirait par conséquent un facteur de décélération entre le test et l'utilisation nominale du composant. D'autant que ce facteur peut devenir significatif si la fréquence de fonctionnement nominale est importante. Pour cela, ce système n'a pas été retenu pour le vieillissement de composants fortement submicroniques CMOS dont la fréquence nominale dépasse bien souvent 100 MHz.

Enfin, les systèmes utilisant la convection thermique localisée ne sont pas non plus adaptés à ce genre de test en raison des diamètres des trompes trop importants (de l'ordre de 10 cm dans le meilleur des cas). Pour ces valeurs, les contraintes thermiques pourraient engendrer une dégradation du boitier ou/et des broches d'assemblage du composant.

Pour toutes les raisons limitatives évoquées précédemment, un système différent a dû être développé. Dans le dispositif utilisé, la température est générée à l'aide d'un module semiconducteur fonctionnant selon l'effet thermoélectrique *Peltier* (Figure 4.25). Ce principe traduit un transfert caloporteur de part et autre des faces du module lorsque celui-ci est alimenté. Lorsque les jonctions p-n du module sont polarisées, le déplacement d'électrons issus des régions n engendre une variation de calories positive au niveau de la jonction et corollairement un déplacement calorifique négatif par l'intermédiaire des trous (Figure 4.26). Ceci se traduit par une différence de température Δ T entre la face supérieure (T_{SUP}) et inférieure (T_{INF}). Pour inverser le flux caloporteur, il suffit alors d'inverser le sens de la polarisation. La face chaude devient froide et la face froide devient chaude.



Figure 4.25: Module Peltier



La relation reliant la tension d'alimentation du module, U, avec la différence de température ΔT est la suivante :

$$U = n \cdot [\alpha'' \cdot \Delta T + R \cdot I] \tag{4.1}$$

avec n le nombre de couples Peltier contenus dans le module, d'une constante exprimée en V/°C qui découle du coefficient du Seebeck du matériau, S, selon la relation S = $n.\alpha$ ", R la résistance électrique par couple (Ω /couple) et I le courant d'alimentation du module (A).

Dans le cadre de notre étude, la face inférieure est utilisée pour vieillir le composant en température. Or, en considérant que le couple U-I du module impose la différence de température Δ T, il devient alors nécessaire de réguler la température sur la face supérieure pour pouvoir imposer une température sur la face inférieure donnée par (4.2). Ceci est d'autant plus vrai s'il s'agit d'appliquer une température plus basse à l'aide de la face inférieure. Pour ce faire, un système de refroidissement à eau comprenant un *waterblock*, un réservoir, deux radiateurs et une pompe, a été utilisé.

$$T_{INF} = T_{SUP} - \frac{1}{\alpha''} \cdot \left(\frac{U}{n} - R \cdot I\right)$$
(4.2)

La conduction thermique entre la face inférieure du module et le composant est assurée à l'aide d'un doigt de cuivre (Figure 4.27). De plus, dans les configurations où la température de la face inférieure serait en dessous à la température de rosée, l'ensemble du dispositif a été moulé dans un sarcophage en téflon de façon à éviter la formation de givre sur le doigt thermique et sur le composant (Figure 4.28). De cette manière, le dispositif permet d'appliquer des températures de l'ordre de 150°C en haute température et jusqu'à -22°C en basse température.

Enfin, les composants SRAM étudiés ont été ouverts mécaniquement en face arrière de manière à pouvoir accéder à la totalité du substrat de la puce. Par ailleurs, l'embout du doigt thermique a été spécialement réalisé pour couvrir l'intégralité de la surface du circuit intégré.





Figure 4.27: Système de contrainte thermique locale en configuration haute température

Figure 4.28: Système de contrainte thermique locale en configuration basse température

L'intérêt d'un tel système est double. Premièrement, il permet d'imposer une température de jonction sur la surface de la puce et non une température ambiante. Ceci présente l'avantage de ne plus tenir compte de la résistance thermique du composant et de pouvoir tester de façon standard, aussi bien les composants à boitiers plastique que ceux en boitiers céramiques. Deuxièmement, le fait que la contrainte soit appliquée localement permet de préserver l'intégrité du testeur, des cartes de tests mais également du boitier et de l'assemblage du composant. De ce fait, il est possible de s'assurer que les dégradations induites par ce dispositif se situent bien au niveau de la puce et non pas à un autre endroit.

4.2.3.3.2 Vieillissement HCI

La combinaison des aspects électriques et thermiques basse température ont permis de réaliser des vieillissements de type HCI accélérés sur deux composants SRAM. Le protocole de test est décrit dans la section ci-dessous.

4.2.3.3.2.1 Conditions de test

La section 4.2.1.1 a décrit les opérations d'écritures successives comme le pire cas d'utilisation d'une cellule mémoire SRAM, vis-à-vis du mécanisme HCI. Toutefois, pour minimiser la durée de l'expérience, il est impératif de réduire au maximum le délai entre deux écritures d'un même point mémoire. Or, puisque les temps d'accès et d'écriture de la cellule sont constants, il est nécessaire de diminuer le nombre de cellules de tests pour reboucler le plus rapidement sur la même cellule sans cumuler un délai trop important lors de la séquence mesure-stress-mesure. De cette manière, sur 4 millions de cellules que comporte la mémoire, seules 32 seront adressées pour le test de vieillissement. Ce choix se traduit par un facteur d'accélération et donc par un gain de temps d'expérience de l'ordre de 125 000 (4M/32).

Le protocole de test effectué sur les deux mémoires de tests est décrit selon la Figure 4.29. Après une première phase de polarisation, la première mémoire (notée A), a été portée à une température de jonction de -10°C tandis que la deuxième (notée B) a été laissée à une température ambiante de 27°C.







Figure 4.30: Banc de vieillissement HCI développé chez EADS IW

Ensuite, une première séquence de 200 millions d'écriture (~72 secondes) a été appliquée en utilisant de manière alternée, un motif de test de 32 bits (x55AA55AA en hexadécimal) ainsi que

son contraire (xAA55AA55). En fin de séquence, un test fonctionnel ainsi qu'une mesure du courant en lecture sont pratiqués de façon à identifier les cellules potentiellement altérées. Le critère de défaillance retenu pour ce test, est soit une adresse défaillante, soit une consommation correspondante à un courant supérieur à la compliance fixée à 40 mA.

Si aucune adresse ne se révèle être défaillante après vérification (lecture) alors le test se poursuit en enchainant les séquences d'écriture décrites plus haut. En cas de défaillance, la mémoire est retestée de façon à bien distinguer les erreurs momentanées des erreurs irréversibles. Enfin, si une défaillance irréversible intervient, alors le test s'arrête en prenant soin de relever l'adresse erronée, le motif ainsi que la consommation.

4.2.3.3.2.2 Résultats et interprétation

Le composant noté A a subi une défaillance irréversible après 1 003 heures de test (équivalent à 10¹³ opérations d'écritures) alors que le composant B est toujours fonctionnel (après 1 003 h de test). La défaillance s'est produite alors que 16 bits sur 32 présentaient simultanément une erreur fonctionnelle et que les autres cellules du composant n'étaient pas altérées. La consommation en lecture mesurée pour le composant A au cours du test de vieillissement, est présentée dans la Figure 4.31. Le tracé montre explicitement 3 zones.

La première zone montre que sur les 870 premières heures du test, le courant est assez proche de sa valeur d'origine. Ceci ne signifie pas qu'il n'y ait pas de dégradation au niveau de l'oxyde de grille des transistors NMOS mais simplement que celle-ci n'a pas encore d'impact sur la consommation générale du composant. La deuxième zone montre une augmentation de la consommation en lecture de l'ordre de 8 %. Cette augmentation n'est toutefois pas assez significative pour générer une erreur fonctionnelle mais peut être interprétée comme un précurseur à la défaillance qui intervient ensuite dans la zone 3. Dans cette zone, la consommation augmente soudainement au-delà de 50 % de sa valeur d'origine jusqu'à ce que le test soit arrêté après une erreur fonctionnelle.

Le composant B ne montre aucune augmentation de la consommation pendant la phase de test. Ce résultat confirme en particulier le fort impact des basses températures sur la dégradation HCI au niveau du produit comme nous l'avons établi au chapitre 2 (section 2.4.2.1.1) au niveau des mécanismes impliqués sur les dispositifs élémentaires (transistors NMOS et PMOS et l'inverseur CMOS).



Figure 4.31: Caractéristique du courant de consommation en phase de lecture pour le composant A au cours du vieillissement HCI

Il reste néanmoins un point à expliciter. Comment la consommation d'une mémoire de 4 Mbits peut varier de plus de 50 % avec seulement 32 cellules potentiellement dégradées ? Cet aspect peut être expliqué par la dégradation progressive de la tension de seuil des transistors NMOS de décharge (N₁ et N₂ sur la Figure 4.8). Cette variation induit un retard de fermeture d'un des NMOS, si bien qu'au bout d'un moment, il est envisageable que le transistor PMOS de décharge reste passant alors que le transistor NMOS ne soit pas encore refermé. Ceci aurait pour conséquence de créer un court-circuit entre les lignes d'alimentation de la cellule V_{DD} et V_{ss} et ainsi une brusque augmentation du courant de consommation.

D'un autre coté, la variation de consommation pourrait également être attribuée à un claquage doux dans la région 2 (Figure 4.31), suivi ensuite d'un claquage franc dans la région 3. Cette détérioration de l'oxyde de grille correspondrait alors au mécanisme CHC décrit dans le chapitre 2, dans la section 2.4.2.1.2. Ce mécanisme traduit la dégradation de l'oxyde de grille localement par les porteurs chauds issus du canal, qui sous l'effet du champ vertical présent dans la région grille-drain, peuvent passer la barrière de potentiel de l'oxyde et ainsi générer des pièges à l'interface, jusqu'au claquage de celui-ci. Cependant ce comportement est surtout observé à haute température, ce qui exclut cette hypothèse étant donné les conditions de tests à basse température opérées sur le composant A.

4.2.3.3.2.3 Discussion

Le temps à défaillance obtenu lors du test précédent pour le composant A, constitue une donnée brute. Or, à partir de ce temps à défaillance noté TTF_{TEST_SRAM} , il est possible de déterminer le temps à défaillance de ce composant s'il avait été utilisé dans une application réelle TTF_{APPLI_SRAM} . En prenant l'exemple des profils de mission 1 et 2 décrites dans le chapitre 3, les temps à défaillance de la SRAM dans les deux applications ont pu être déterminés en appliquant un facteur d'accélération dit « produit ». Ce facteur reliant TTF_{TEST_SRAM} à TTF_{APPLI_SRAM} est donné par l'expression (4.3) :

$$TTF_{APPLI_SRAM} = TTF_{TEST_SRAM} \cdot \left(\frac{1}{\alpha}\right) \cdot \frac{NB_{MEM}}{NB_{TEST}} \cdot \frac{1}{t_{ON}} \cdot \sum_{i}^{N} \left(e^{\frac{E_A}{k} \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right)} \cdot t_i\right)$$
(4.3)

avec α le pourcentage de sollicitation des transistors lorsque l'équipement est sous tension (fixé à 1), NB_{MEM} la densité de la mémoire, NB_{TEST} le nombre de cellules testées, t_{ON} le temps pendant lequel l'équipement est en fonctionnement (en heures), N le nombre de plateaux de température au cours de la mission, E_A l'énergie d'activation du mécanisme correspondant à la technologie de la SRAM (= 0,2 eV), k la constante de Boltzmann, t_i le temps pendant lequel l'équipement fonctionne à la température T_{APPLI} et enfin T_{EST} la température de test.

D'après l'équation (4.3), le facteur d'accélération appliqué pour le profil 1 est estimé à $5,5.10^5$ et à $5,1.10^5$ pour le profil 2. Par exemple, cela signifie que 1 seconde de test pratiquée sur 32 cellules de la mémoire SRAM à -10°C est équivalent à plus de 6 mois d'utilisation de l'ensemble du plan mémoire dans les conditions d'utilisation définies par le profil 1.

Il devient maintenant intéressant de comparer les temps de défaillances de la mémoire en application TTF_{APPLI_SRAM} avec ceux calculés à l'aide de la méthode physique proposée dans le chapitre 3, qui sont notés $TTF_{APPLI_PHYSIQUE}$. Pour ce faire, le modèle de dégradation PTM HCI présenté dans l'équation (2.35) a été utilisé pour calculer $TTF_{APPLI_PHYSIQUE}$ pour le nœud 90nm, avec pour entrées : les paramètres technologiques de la mémoire présentés dans la section 4.2.2, ainsi que la tension de seuil V_{TH} extraite dans la section 4.2.3.1.2. Les nœuds 130, 65 et 45 nm *Basse consommation* ont également été représentés à titre de perspectives. Pour ces nœuds, les données d'entrées sont issues de la rubrique basse consommation de l'ITRS **[ITR 07]**. Les résultats sont présentés dans la Figure 4.32 pour le profil 1 et dans la Figure 4.33 pour le profil 2.



Figure 4.32: Comparaison pour le profil 1,entre les prédictions de durée de vie HCI issues de la méthodologie entre les nœuds 130 et 45 nm *Basse consommation* (critère $V_{TH} = 10\%$), et la défaillance HCI de la SRAM 90 nm LP



Figure 4.33: Comparaison pour le profil 2,entre les prédictions de durée de vie HCI issues de la méthodologie entre les nœuds 130 et 45 nm *Basse consommation* (critère $V_{TH} = 10\%$), et la défaillance HCI de la SRAM 90 nm LP

Les résultats obtenus pour les 2 profils, montrent une assez bonne cohérence des TTF_{APPLI_SRAM} avec les prédictions faites au niveau du silicium $TTF_{APPLI_PHYSIQUE}$. En ce qui concerne le nœud 90 nm, l'écart relatif calculé pour l'application 2 est de l'ordre de 32 % alors que celui estimée pour le profil 1 se situe aux alentours de 7 %.

Bien entendu pour valider cette approche, il est nécessaire de procéder à des vieillissements sur une population de composant plus importante. Ceci permettrait d'une part de tenir compte de la variabilité des procédés de fabrication et d'autre part de resituer le temps à défaillance obtenu TTF_{TEST_SRAM} par rapport à une distribution de défaillance. Cependant, les résultats sur le 90 nm sont assez encourageants pour ces utilisations de type avionique.

Les temps de défaillance en application TTF_{APPLL_SRAM} obtenus pour la SRAM sont très importants (de l'ordre de 10⁵ années) et bien aux dessus des spécifications. Ceci confirme que le mécanisme HCI n'est pas problématique pour les deux applications aéronautiques, concernant les composants de génération 90 nm *basse consommation* disposant d'oxydes de grille de transistors nitrurés SiON. Néanmoins, selon l'extrapolation des lois d'accélérations, ils pourraient le devenir en dessous des nœuds 32 nm LP pour le profil 1 (Figure 4.32) et 20 nm LP pour le profil 2 (Figure 4.33).

4.2.3.3.3 Vieillissement NBTI

De la même façon que pour le mécanisme HCI, le testeur de composant ELECTRE ainsi que le système de contrainte thermique locale ont été utilisés pour effectuer des vieillissements de type NBTI sur des composants SRAM.

4.2.3.3.3.1 Conditions de test

Le stockage à haute température d'une même donnée au sein d'une cellule 6T, a été identifié comme une utilisation pire cas NBTI, dans les transistors PMOS de la SRAM. Ceci indique que la fréquence d'adressage n'est pas un facteur d'accélération du mécanisme. De ce fait, la totalité des cellules a été adressée pour ce test (4 Mbits).

Le protocole de test décrit dans la Figure 4.34 a été appliqué sur une mémoire contrainte à une température de jonction de 100°C, sur une durée de 2 200 heures.



Figure 4.34: Protocole de test pour le vieillissement NBTI sur SRAM (100°C)

Après la mise en température, les 4 millions de cellules SRAM que comporte la mémoire ont été écrites à l'aide d'un même motif de test (x55AA55AA en hexadécimal). Les données sont ensuite stockées pendant la durée d'une heure. A la fin de cette période, une mesure de consommation en absence de commande (IDLE state) est effectuée ainsi qu'un test de lecture de façon à identifier une cellule potentiellement défaillante. Comme pour le mécanisme HCI, le critère de défaillance retenu, est soit une adresse défaillante, soit une consommation supérieure à un courant de compliance de 40 mA. Si aucune erreur fonctionnelle n'a été repérée alors, le plan mémoire est réécrit à l'aide du même motif de façon à ne pas relaxer le mécanisme et la procédure se répète. En cas de défaillance, la mémoire est réinitialisée avant d'être testée une nouvelle fois. Si la défaillance est toujours présente, alors le test de vieillissement se termine en relevant l'adresse erronée, le motif ainsi que la consommation.

4.2.3.3.3.2 Résultats et interprétations

Le test de vieillissement n'a pas mis en évidence une défaillance de composant comme cela a pu être le cas pour le test HCI. Le courant de consommation en lecture ainsi que le courant en absence de commande (IDLE state) ont pu être toutefois mesurés au cours des 2 200 heures de test (Figure 4.35 et Figure 4.36). On observe sur les deux tracés présentés qu'aucune dérive brutale de consommation n'a pu être constatée mais plutôt une croissance lente et progressive des courants. Cependant, il apparaît deux légères augmentations de courant plus nettes, respectivement au bout de 440 et 1 050 heures de test. Ces augmentations sont de l'ordre de 3% et coïncident avec les moments où le banc de test NBTI à du être redémarré à la suite à de problèmes techniques.

Ces rehaussements de consommation peuvent être dus directement à l'effet de relaxation (post contrainte) des transistors PMOS quand la tension de fonctionnement est enlevée, ce qui à la remise sous tension permet aux blocs de charges (PMOS) de la cellule de fournir plus de courant que celui atteint sous contrainte d'usure lente, ce qui est propre au mécanisme NBTI. En toute rigueur, la valeur de la pente de la dégradation en loi de puissance du temps, extraite sur la Figure 4.36, en l'occurrence 0,02, est 8 fois inférieure à celles mesurées classiquement sur des transistors PMOS soumis à une contrainte NBTI. De ce fait, ces augmentations lentes de consommation ne peuvent être attribuées à une dégradation NBTI.

Chapitre 4 : Démonstration, validation expérimentale et méthodologique pour la sélection de composants





Figure 4.35: Evolution des courants de lecture (en rouge) et d'absence de commande IDL (en bleu) au cours du test de vieillissement NBTI, SRAM 90 nm, oxyde SiON

Figure 4.36: Evolution du courant d'absence de commande au cours du test de vieillissement NBTI, SRAM 90 nm, oxyde SiON

4.2.3.3.3.3 Discussions

Le fait qu'aucune défaillance ne soit intervenue au cours du test ne signifie pas pour autant que le mécanisme NBTI ne soit pas un risque significatif pour la SRAM 90 nm. Ceci signifie simplement que les conditions de vieillissement ne sont pas assez sévères pour provoquer une défaillance dans un temps raisonnable, en d'autres termes, le facteur d'accélération en température et tension n'est pas assez important. En effet, en utilisant le facteur décrit par l'équation (4.4) avec une énergie d'activation de 0,18 eV correspondant aux oxydes de grille SiON d'épaisseur 2 nm **[TSU 05]**, il est possible de calculer l'accélération du mécanisme à partir de l'expression (4.4) :

$$TTF_{APPLI_SRAM} = TTF_{TEST_SRAM} \cdot \left(\frac{\alpha}{100}\right) \cdot \frac{1}{t_{ON}} \cdot \sum_{i}^{N} \left(e^{\frac{E_A}{k} \left(\frac{1}{T_{APPLI}} - \frac{1}{T_{TEST}}\right)} \cdot t_i\right)$$
(4.4)

Pour le profil 1 par exemple, 1 heure de test NBTI à 100°C équivaut à seulement 2,8 heures de l'application réelle si les transistors PMOS de la mémoire sont sollicités de manière statique ($\alpha = 100 \%$) et à 280 heures s'ils le sont pendant 1 % du temps de la mission.

Or, pour connaitre l'ordre de grandeur du temps à défaillance, les calculs de prédiction de durée de vie ont été effectués pour les technologies d'oxyde SiON de génération 130 nm à 45 nm *basse consommation,* (Figure 4.37). Selon ces estimations, le temps de défaillance de la SRAM 90 nm vis à vis du mécanisme NBTI serait de l'ordre de 13 000 années (Figure 4.37). C'est-à-dire que

selon l'expression (4.4), il serait nécessaire d'effectuer plus de 4 600 années de test pour provoquer une défaillance si $\alpha = 100$ % et 46 ans si $\alpha = 1$ %. Ceci signifie que le temps à défaillance sous contrainte NBTI est très difficilement atteignable pour cette mémoire à ce facteur d'accélération fixé, même si la contrainte thermique appliquée au cours du test était plus importante. Le temps à la défaillance reste bien au-delà des spécifications des profils de mission.



Figure 4.37: Prédictions de durée de vie de technologies SiON 130, 90, 65 et 45 nm basse consommation, soumises au profil de mission 1, application de la méthode décrite dans le chapitre 3 avec pour entrée les données de l'ITRS LP **[ITR 07]**

Cependant, l'évaluation du bon fonctionnement d'un composant pendant une période spécifiée ne nécessite pas nécessairement de connaitre le temps à défaillance du composant. Il est en effet possible de procéder au raisonnement inverse et de calculer le nombre d'heures de test nécessaires à effectuer, de manière à atteindre la spécification en durée de vie du composant. Selon l'expression (4.4) appliquée au profil 1, les 30 ans d'utilisation seraient couverts par 74 250 heures de test NBTI à 100°C si $\alpha = 100$ % et à seulement 742 heures si $\alpha = 1$ %. Dans ce dernier cas, les 2 200 heures de test effectuées sans qu'aucune défaillance n'ait été observée permet de conclure à la non-criticité du mécanisme NBTI pour ces cellules SRAM 90nm.

De même, en tenant compte de la durée maximum de test que nous avons atteint, soit 2 200 heures, il est possible de calculer le facteur de sollicitation α approprié de façon encore une fois, à couvrir la durée de fonctionnement. Dans ces conditions, il est possible de conclure que le mécanisme NBTI ne représente pas un risque de défaillance majeur pour cette mémoire, si les transistors PMOS ne sont pas sollicités plus de 33 % du temps pendant la mission.

4.2.4 Bilan

La faisabilité de plusieurs techniques de vieillissement de composant fortement submicroniques a été étudiée et appliquée sur des mémoires SRAM de génération 90 nm. Parmi ces techniques, le banc de test à l'aide de nano pointes de tungstène à permis de tracer des caractéristiques I_{DS} - V_{GS} et I_{DS} - V_{DS} de transistor NMOS 90 nm et d'en extraire les paramètres électriques tels que la tension de seuil et le courant de drain en saturation. Néanmoins, ce banc n'a pas pu être adapté pour réaliser des vieillissements de composant en raison de la faible durée de vie des nano pointes. Il en est de même en ce qui concerne le banc de test utilisant des plots de contact FIB. La grande variabilité introduite par la préparation de l'échantillon n'a pas permis d'obtenir des résultats cohérents. De ce fait, cette technique n'est pas recommandée pour réaliser des vieillissements de composant.

En revanche, le banc de test constitué du testeur de composant ELECTRE ainsi que du système de contrainte thermique local développé au cours de ces travaux, semble plus prometteur. En effet, la possibilité d'appliquer une température sur la surface de la puce associé aux instructions pire cas générées par le testeur, ont permis de réaliser des vieillissements de type NBTI et HCI sur SRAM. Les résultats obtenus au cours de ces tests on montré que les mécanismes de défaillances HCI et NBTI ne représentaient pas de risque majeur pour la mémoire SRAM si elle était utilisée dans le cadre de l'application 1. D'autre part, le temps à défaillance obtenu dans le cadre du test HCI, est relativement proche des prédictions issues de la méthodologie développée dans le chapitre 3, pour le nœud 90 nm basse consommation. Bien entendu, pour valider cette approche, il serait nécessaire de reproduire le test non plus sur un seul composant mais sur une population plus importante. C'est pourquoi un testeur permettant de vieillir une douzaine de composant en parallèle lors d'un même test, est en cours de développement chez EADS IW. De plus, la validité des prédictions réalisées pour les autres nœuds technologiques doit être validée par des tests sur des composants de même type mais de générations différentes.

Les tests expérimentaux permettent de quantifier le temps à défaillance des composants dans une application donnée. Cependant, cette approche est relativement lourde à mettre en place lorsqu'il s'agit d'évaluer la fiabilité de plusieurs composants sur un équipement. Pour cela, une méthodologie de sélection de composant par la technologie a été développée en se fondant d'une part sur le profil de mission et les paramètres technologiques, ainsi que sur les résultats de tests fabricants pratiqués au niveau du produit.

4.3 Sélection de composant par les données de test fabricant

Le temps à défaillance d'un composant fortement submicronique peut être estimé au niveau silicium à l'aide des modèles de dégradation présentés dans le chapitre 2 et validés au niveau du produit à l'aide des tests expérimentaux présentés en début de chapitre 4. Toutefois, cette approche peut s'avérer périlleuse si le temps à défaillance du composant demeure très grand comparé à la durée d'utilisation de l'équipement. C'est pourquoi, plutôt que de chercher à connaitre le temps de défaillance d'un composant dans une application (ITF), il devient préférable d'estimer le temps minium de bon fonctionnement (TMBF). Le temps minium de bon fonctionnement du composant peut être obtenu en interprétant les résultats de tests pratiqués par les fabricants. En ce sens, la section suivante propose une méthode de sélection composant au niveau du produit en fonction du profil de mission, de la technologie et des résultats de tests pratiqués par les fabricants.

4.3.1 Méthode

4.3.1.1 Technologie du composant

La première étape de cette méthodologie est de collecter toutes les informations relatives à la technologie du composant étudié, à savoir : le fabricant, la fonderie où le circuit intégré a été fabriqué, les données technologiques BEOL et FEOL. Les données technologiques sont rarement accessibles pour un utilisateur, c'est pourquoi il est bien souvent nécessaire de procéder à des analyses de construction du circuit intégré, de façon à extraire ces paramètres. Les données technologiques ciblées sont :

- Le pas lithographique ou génération technologique,
- La nature ainsi que les dimensions des lignes de métallisation BEOL,
- Les matériaux, les dimensions des transistors FEOL,
- La nature du substrat du circuit intégré (silicium massique, SOI, silicium étiré, GaAs, structures 3D etc...)

L'objectif est de pouvoir établir une carte d'identité technologique pour chaque référence de composant étudiée. Dans cette optique, une base de données *technologique* des composants étudiés au cours de ces travaux, a été créée.

4.3.1.2 Profil de mission

La seconde étape consiste à décrire le plus fidèlement possible, l'environnement dans lequel est sollicité le composant. En toute rigueur, une analyse de risque complète d'un composant nécessite de connaitre tous les facteurs qui peuvent potentiellement impacter la fiabilité de l'assemblage, du boitier d'encapsulation ainsi que de la puce d'un composant. Toutefois, dans le cadre de cette étude, nous ne considérerons que ceux liés à la dégradation du circuit intégré, en l'occurrence les facteurs électriques et thermiques. De ce fait, les données recherchées sont :

- Le cycle de fonctionnement journalier du composant,
- Le profil de mission en température,
- Le nombre de jour d'utilisation par années,
- Le nombre d'année d'utilisation.

4.3.1.3 Matrice de risques

La connaissance conjointe de la technologie du composant ainsi que des conditions d'utilisation permettent de dresser une matrice de risques spécifiques du composant. En associant une criticité à un mécanisme de dégradation, cette analyse permet d'éliminer les mécanismes non pertinents et corolairement de ne retenir que ceux qui sont susceptibles de se produire. Pour cela, il est nécessaire de bien connaitre d'une part, les mécanismes de dégradation ainsi que leur catalyseur environnementaux et d'autre part, la technologie et le fonctionnement des composants.

L'analyse doit être menée selon des critères d'utilisation. Par exemple, les mécanismes d'usure thermoélectriques tels que l'électromigration ou encore les porteurs chauds peuvent être désignés comme étant non pertinents pour un composant utilisé dans des applications de stockage non alimenté. De plus, l'évaluation de la criticité doit également prendre en considération la fonctionnalité électrique du composant. Par exemple, les mécanismes FEOL d'instabilité en température et de claquage d'oxyde ne peuvent se produire sur des transistors de type flash NOR/NAND ou sur des structures DRAM. La principale raison est qu'il n'existe aucune contrainte électrique statique au niveau de l'oxyde de grille de ces dispositifs. De ce fait, ces mécanismes sont également jugés non pertinents pour ces structures.

Enfin, la connaissance de la technologie et des matériaux utilisés dans le circuit intégré peuvent renseigner sur des risques potentiels de fiabilité. Le chapitre 3 a notamment montré par exemple, le fort impact que pouvaient avoir les oxydes de grille nitrurés sur la durée de vie d'un composant utilisé pour une application avionique.

4.3.1.4 Calcul du taux et du temps minimum de bon fonctionnement du composant déterminés dans les conditions de test

L'élimination des mécanismes non pertinents permet de recentrer l'étude sur les mécanismes d'endommagement potentiels. Aussi, la prochaine étape consiste à extraire dans les rapports de qualification des fabricants, les tests de dégradation pertinents pouvant conduire à l'apparition des mécanismes identifiés. Dans cette optique, le Tableau 4.1 présente les différents tests pratiqués au niveau du produit, susceptibles de provoquer les mécanismes de défaillance du silicium. Les conditions opératoires des tests mentionnés peuvent varier en fonction des normes, Ceci est notamment le cas du test d'endurance HTOL *High Temperature Operating Life* où la température de test spécifiée peuvent varier de 125°C jusqu'à 150°C selon les normes.

Parfois certains de ces tests ne font référence à aucune norme pour les produits commerciaux et sont pratiqués dans des conditions différentes selon les fabricants (nombre de composant, durée de test, critères de défaillances etc...). A l'inverse, certains grands fabricants (IBM, STMicroelectronics, TSMC, Toshiba...) essaient de se regrouper pour faire admettre des normes de tests au niveau silicium comme avec les normes JEDEC (JC-14) ou NIST [JED 11].

Si les résultats de tests ne sont pas accessibles, cette étape permet d'indiquer à un utilisateur les tests qu'il serait nécessaire d'effectuer pour s'assurer de la fiabilité du composant.

Zones du	Mécanismes de défaillance	Tests produit	Exemple de norme
circuit integre			
	Electromigration	Endurance HTOL	JESD22-A104 JESD22-A108
BEOL	Stress Migration (fatigue des barrières de diffusion)	Cycles thermiques	Non normé
	TDDB inter métallique	Endurance HTOL	MIL-STD-883 JESD22-A104 JESD22-A108
	TDDB oxyde de grille	Test de stockage de données sur zone mémoire	Non normé
FEOL	Porteurs chauds	Life Test Dynamique	Non normé
	Instabilité en température	Test de stockage de données sur zone mémoire	Non normé

Tableau 4.1: Correspondance entre les mécanismes de défaillance du silicium et les tests fabricants opérés au niveau du produit

L'analyse de risque fournie également une indication quant à la nature des endommagements potentiels, progressifs ou brusques. Il est en effet possible de distinguer les défaillances dites aléatoires de celles qui sont caractéristiques d'une dégradation d'usure. La répartition des différents mécanismes a été synthétisée sur une courbe en baignoire qui est proposée Figure 4.38. Pour des mécanismes aléatoires, le paramètre de fiabilité recherché sera le taux de défaillance en application (λ_{APPLI}) alors que dans le cas des mécanismes d'usure, le paramètre ciblé sera le temps minimum de bon fonctionnement en application (TMBF).



Figure 4.38 : Schématique de la répartition des différents mécanismes de défaillance du silicium sur une représentation du taux de défaillance instantané.

4.3.1.4.1 Calcul du taux de défaillance

$$\lambda_{TEST} = \frac{\chi_{1-\alpha'''}^2(2c+2)}{2 \cdot t^*} \cdot 10^{-9}$$
(4.5)

Avec t^{*} le nombre d'heure-composant exprimé par (4.6) où t_i est le temps au bout duquel un composant i est devenu défaillant.

$$t^* = N \cdot t_{TEST} - \sum_{i=0}^{c} i \cdot t_i \tag{4.6}$$

4.3.1.4.2 Calcul du temps minimum de bon fonctionnement TMBF

Pour les mécanismes d'usure, il existe deux cas de figure. Dans le cas où les tests du fabricant permettent de révéler des défaillances d'usure, alors le temps moyen à défaillance du composant peut être extrait et comparé à la spécification en durée de vie de l'application après avoir préalablement appliqué le facteur d'accélération qui convient.

Si en revanche les tests fabricants ne permettent pas de mettre en évidence des défaillances d'usure, alors il devient nécessaire d'utiliser la notion de temps minimum de bon fonctionnement (TMBF) au lieu du temps à défaillance du composant (TTF).

Le point de départ du raisonnement provient du fait qu'une défaillance de composant intervient toujours, soit par le mécanisme aléatoire le plus probable, soit par le mécanisme d'usure le plus actif. De ce fait, la probabilité de défaillance F(t) d'une population N de composant peut s'exprimer à l'aide d'une loi bimodale (équation (4.8)). La première partie de cette loi est décrite par un formalisme de type exponentiel, où λ représente le taux de défaillance du mécanisme aléatoire le plus probable. La deuxième partie est elle représentative du mécanisme d'usure le plus prépondérant, c'est-à-dire celui dont le temps à défaillance moyen (TTF_{63,2%}) est le plus petit (Figure 4.39).

 $F(t) = \begin{cases} 1 - e^{-\lambda \cdot t} & \text{Pour les mécanismes aléatoires} \\ 1 - e^{-\left(\frac{t}{TTF_{63,2\%}}\right)^{\beta}} & \text{Pour les mécanismes d'usure} \end{cases}$ (4.7)



Figure 4.39: Représentation schématique du pourcentage de défaillance d'une population de N composants

La continuité de F(t) est vraie quel que soit t (Figure 4.39). Cependant, ce point d'inflexion peut s'avérer périlleux à atteindre si les durées de tests pratiqués par les fabricants, ne sont pas assez longues ou encore, si les conditions de ces tests ne sont pas assez sévères. C'est pourquoi, plutôt que raisonner sur le temps de défaillance moyen du composant, il devient alors plus intéressant d'étudier le temps minimum de bon fonctionnement garanti par les tests du fabricant. Pour cela, il est nécessaire de considérer un pire cas dans lequel la probabilité de défaillance par mécanisme d'usure prend le pas sur la probabilité de défaillance aléatoire à l'issue du test fabricant. Ceci signifie que le point de continuité préalablement décrit, interviendrait pour t = t_{TEST} tel qu'il est décrit dans la Figure 4.40 (point 1).



Figure 4.40: Représentation schématique du temps minimum de bon fonctionnement TMBF à l'issu de test fabricant

De ce fait, l'équation (4.7) peut être réécrite en :

$$F(t) = \begin{cases} 1 - e^{-\lambda_{TEST} \cdot t_{TEST}} & \text{si } t_{0\%} < t \le t_{TEST} \\ 1 - e^{-\left(\frac{t}{TMBF_{TEST_{63,2\%}}}\right)^{\beta}} & \text{si } t_{TEST} \le t < t_{100\%} \end{cases}$$
(4.8)

où λ_{TEST} est le taux de défaillance du composant issu des tests avec c+1 pièces défaillantes, où TMBF_{TEST_63,2%} est le temps minium de bon fonctionnement moyen et β la pente du mécanisme d'usure le plus prépondérant, dépendante de la technologie du composant (Figure 4.40). Dans les conditions de test, il est alors possible d'écrire par continuité les expressions suivantes:

$$TMBF_{TEST_F(t_{TEST})} = t_{TEST}$$
 Point 1 sur la Figure 4.40 (4.9)

$$TMBF_{TEST_{63,2\%}} = \frac{(t_{TEST})^{1-\frac{1}{\beta}}}{\lambda_{TEST}^{\frac{1}{\beta}}} \qquad \text{Point 2 sur la Figure 4.40}$$
(4.10)

Le temps minium de bon fonctionnement est par défaut calculé pour 63,2 % d'échantillons défaillants, ce qui constitue une valeur moyenne dans le formalisme de Weibull. Toutefois, d'autres critères de population peuvent être choisis en fonction du besoin de fiabilité d'une application. Pour cela, la relation reliant le pourcentage d'échantillon défaillant F au temps minimum de bon fonctionnement est donné par l'expression (4.11) :

$$TMBF_{TEST_F\%} = TMBF_{TEST_{63,2\%}} \cdot [-ln(1-F)]^{\frac{1}{\beta}}$$
 (4.11)

Cette démarche est à réaliser pour tous les tests jugés d'intérêt. Dans le cas où un test pourrait provoquer deux mécanismes pertinents, alors la criticité des mécanismes dans les conditions d'application sera déterminée à l'aide de leurs facteurs d'accélérations respectifs.

4.3.1.5 Facteur d'accélération (AF)

Le facteur d'accélération est une modélisation mathématique d'un phénomène de dégradation physique. Il est utilisé pour déterminer la valeur d'un paramètre de fiabilité d'un composant sur des longues durées et sous faible contrainte ou aux tensions de fonctionnement, à partir d'essais accélérés opérés sur des courtes durées et à forte contraintes (tension, température, courant). Aussi, la bonne estimation du taux de défaillance ainsi que du temps minimum de bon fonctionnement d'un composant, dépendent de la précision du facteur d'accélération qui est appliqué.

Pour les mécanismes aléatoires, un facteur d'accélération de type Arrhenius est classiquement utilisé pour couvrir toutes les défaillances liées à la température (noté $AF_{aléatoire}$). Toutefois, la valeur de l'énergie d'activation égale à 0,7 eV, couramment utilisée par les fabricants reste somme toute discutable. Cet aspect ne sera pas traité dans ce manuscrit.

En ce qui concerne les mécanismes d'usure, le chapitre 2 a montré qu'au niveau du silicium, il n'existait pas de facteur d'accélération standard mais bien plusieurs lois de dégradation suivant l'importance du mécanisme. Celles-ci dépendent notamment du mécanisme de défaillance étudié et des paramètres BEOL et FEOL du composant. En ce sens, la connaissance des paramètres technologiques du composant (section 4.3.1.1), permet de sélectionner et d'appliquer par mécanisme, le facteur d'accélération le plus pertinent.

4.3.1.6 Calcul du taux et du temps minimum de bon fonctionnement du composant dans les conditions de l'application

Une fois les facteurs d'accélération déterminés, le taux de défaillance d'un composant en application peut être calculé à partir de l'expression (4.12). De même, le temps minimum de bon fonctionnent dans les conditions de l'application est exprimé pour les mécanismes d'usure à l'aide de l'équation (4.13).

$$\lambda_{APPLI} = \frac{\lambda_{TEST}}{AF_{aléatoire}} \tag{4.12}$$

$$TMBF_{APPLI_{63,2\%}} = TMBF_{TEST_{63,2\%}} \cdot AF_{usure}$$

$$(4.13)$$

La méthodologie développée a été synthétisée dans la Figure 4.41. Elle peut être utilisée de deux façons différentes. La première permet d'estimer le temps minium de bon fonctionnement d'une référence commerciale pour une application spécifique. La deuxième permet de sélectionner parmi des composants présentant une fonction identique ainsi que des performances électriques équivalentes, la référence qui présentera le temps minimum de bon fonctionnement le plus important vis-à-vis des spécifications d'une application.

De façon à illustrer ces principes, un cas d'application de la méthodologie est proposé dans la section suivante, dans lequel un utilisateur sélectionne pour son application, une mémoire flash NAND parmi deux références commerciales.





4.3.2 Application de la méthodologie sur deux mémoires flash NAND

4.3.2.1 Technologie des composants flash

Les composants choisis pour cette étude sont deux mémoires flash NAND 32 Gbits de génération 50 nm présentant des performances électriques équivalentes. Pour des raisons de confidentialité, la provenance de ces composants ne sera pas communiquée. La première mémoire est conçue par le fabricant A, fabriquée à l'aide de métallisations en aluminium isolées par de la silice. La seconde mémoire en revanche, est issue d'un procédé de fabrication de type dual damascène en cuivre dont les lignes métalliques sont isolées par un oxyde à basse permittivité (fabricant B). Pour les deux mémoires, la tension nominale est V_{DD} = 3,3 V



Figure 4.42: Coupe micrographique du composant A, microscope électronique à balayage, x 13 000

Figure 4.43: Coupe micrographique du composant B, microscope électronique à balayage, x 13 000

Concernant les aspects FEOL, les deux mémoires sont de technologie à grille flottante dont l'oxyde de grille est constitué de silice d'épaisseur 7 nm. Enfin, le substrat des deux circuits intégrés est constitué de silicium massique. Une synthèse des aspects technologiques des deux mémoires est proposée dans le Tableau 4.2.

Zones	Paramètres technologiques	Fabricant A	Fabricant B
	Génération technologique	50 nm	50 nm
EEOI	Métallisation BEOL	Aluminium	Cuivre
FEOL	Isolation intermétallique	SiO2	Low-K SiOCH
	Via d'interconnexion	tungstène	cuivre
	Structure du point mémoire	Grille flottante	Grille flottante
BEOL	Nature oxyde de grille	SiO ₂ 7 nm	SiO ₂ 7 nm
	Nature du substrat	Massique	Massique

Tableau 4.2: Synthèse des paramètres technologiques des deux mémoires flash, issues respectivement des fabricants A et B.

4.3.2.2 Profil de mission

Le profil de mission appliqué est celui d'un calculateur de commande de vol situé dans une baie avionique. Le profil proposé est celui décrit dans la section 3.2.1.1 du chapitre 3 avec trois phases d'utilisations (stockage, tests, exploitation).

4.3.2.3 Analyse de criticité

Les données technologiques et environnementales ont permis de réaliser une analyse de criticité pour les deux types de mémoires. L'analyse a permis d'identifier deux mécanismes d'usure pertinents, à savoir l'électromigration et les porteurs chauds (Tableau 4.3).

Le mécanisme d'électromigration a été sélectionné dans le cas de la mémoire A pour deux raisons. La première raison provient du fait que les métallisations en aluminium sont jugées intrinsèquement plus sensibles que celles en cuivre. Deuxièmement, les métallisations du dernier niveau BEOL du composant A sont relativement longues (supérieures à 100 μ m, c'est à dire supérieure à la longueur de Blech calculée pour cette ligne) et larges (supérieures à 2 μ m) ce qui favorise la diffusion à la frontière des grains. Toutefois, la température en mission de l'application ne dépasse pas les 55°C, ce qui signifie que l'accélération du mécanisme est assez faible.

Zanas	Daramètres technologiques	Pertinence des mécanismes		
Zones	Farametres technologiques	Composant A	Composant B	
	Electromigration	Moyenne	Faible	
BEOL	Stress Migration	Faible	Faible	
	Claquage de l'oxyde intermétallique	Faible	Faible	
	Claquage de l'oxyde de grille	Nulle	Nulle	
FEOL	Porteurs chauds	Moyenne	Moyenne	
	Instabilités en température	Nulle	Nulle	

Tableau 4.3 : Analyse de criticité des deux mémoires flash

La dégradation par porteurs chauds a également été retenue car le mécanisme HCI est le mode d'injections de porteurs avec lequel les deux mémoires sont écrites. Pour des flash NAND, le mécanisme HCI impose un critère d'évaluation relativement sévère (chapitre 2), il a été appliqué à plusieurs types de fabricants montrant clairement un niveau de criticité non adapté aux missions avioniques et aérospatiales décrites dans le chapitre 3. Ceci a rendu l'analyse et la divulgation des résultats de type produits hautement confidentiels au moment de cette étude, qui ne nous ont pas permis de les détailler dans ce manuscrit. Néanmoins, ceci a confirmé que la méthodologie développée se montre à la fois applicable à différents produits mais aussi permet de distinguer de façon précise les différents facteurs d'accélération et sensibilités des différents constructeurs.

En revanche le mécanisme de stress migration a été écarté en raison du faible écart de température de l'application entre les phases où le calculateur est en fonctionnement et celles où il ne l'est pas (estimé à 40°C). De même, l'espacement relativement important entre les métallisations BEOL conjugué au fait que la température maximale d'utilisation ne dépasse pas les 55°C, permettent d'éliminer le mécanisme de claquage d'oxyde intermétallique. Enfin, les mécanismes de claquage d'oxyde de grille et d'instabilité en température ont été jugés non pertinents car le fonctionnement d'une mémoire flash impose qu'il n'existe aucune contrainte statique au niveau du point mémoire. En effet, le transistor à grille flottante ne s'use que si la donnée contenue en mémoire est effacée ou reprogrammée.

4.3.2.4 Test de fiabilité

Le test retenu pour mettre en évidence le mécanisme d'électromigration est le test HTOL. Le test HTOL a été réalisé sur une population de 87 composants dans le cas du fabricant A et de 96 composants pour le fabricant B. Les conditions de tests appliquées par les deux fabricants sont identiques (Tableau 4.4), à savoir, une polarisation égale à 120 % de la tension d'alimentation nominale V_{DD} et une température de jonction fixée à 125°C.

Mécanisme	Test	Fabricant	Conditions	Nombre de pièces de test	Nombre de pièces défaillantes	λ _{TEST} (FITs)
ЕМ	Life Test HTOL	А	125°C, 120%V _{DD} 1 000 h	87	0	~ 10 000
		В	125°C, 120%V _{DD} 1 000 h	96	0	~ 10 000

Tableau 4.4: Tests de qualifications pertinents concernant les mécanismes d'électromigration, pratiqués par les fabricants A et B

A la fin des 1 000 heures de test, aucune pièce défaillante n'a été enregistrée par les deux fabricants. Enfin, en considérant un pourcentage de confiance de 60%, le taux de défaillance des

deux mémoires dans les conditions de test a été estimé à environ 10 000 FITs dans les conditions de tests. Cette valeur est assez importante. Toutefois, ceci ne signifie pas que la fiabilité intrinsèque de ces composants est mauvaise mais plutôt que le nombre d'heures-composants employé pour ces tests est relativement faible.

4.3.2.5 Calcul du temps minium de bon fonctionnement en test

Le temps minimum de bon fonctionnement peut être calculé dans les conditions de test en utilisant l'équation (4.10) avec pour entrée les taux de défaillance calculés dans la section précédente, la durée des tests ainsi que la pente de la distribution de Weibull représentative d'une défaillance par électromigration. Une pente égale à 6 dans le cas de l'aluminium et de 4,5 pour le cuivre ont été appliquées. Ces valeurs sont issues d'essai fabricant dont la provenance ne peut être communiquée dans ce manuscrit. Les $MTBF_{TEST_{63,2\%}}$ ainsi calculés sont égaux à 2 136 h pour le composant A et à 2 812 h pour le composant B.

4.3.2.6 Facteurs d'accélération

Le facteur d'accélération utilisé dans ce cas d'application est le modèle de Black décrit dans le chapitre 2 par l'équation (2.12). La densité de courant traversant les lignes est elle exprimée par la relation (2.13). De ce fait, en combinant les relations (2.12) et (2.13), le facteur d'endommagement peut être réécrit à l'aide de l'équation (4.14) avec m le nombre de palier de température définis par le profil.

$$AF = \left(\frac{V_{APPLI}}{V_{TEST}}\right)^{-n} \cdot \sum_{i}^{m} e^{\frac{E_A}{k} \cdot \left(\frac{1}{T_{TEST}} - \frac{1}{T_{APPLI_i}}\right)}$$
(4.14)

Les paramètres du modèle correspondant aux technologies aluminium (composant A) et cuivre (composant B) ainsi des valeurs des facteurs d'accélérations sont consignés dans le Tableau 4.5. Les calculs effectués à l'aide de (4.14) tiennent compte en entrée, de la variabilité de l'énergie d'activation en température du mécanisme, ce qui se traduit par une gamme de facteur d'accélération en sortie.

Composant	n	E _A	AF
Α	2,02 [BIE 04]	0,5 - 0,6 [JEP 11]	67 - 145
В	1,40 [LIN 07]	0,8 - 0,9 [GUI 05]	620 - 1350

Tableau 4.5: Valeur des paramètres du modèle de dégradation et des facteurs d'accélération associés

4.3.2.7 Calcul du temps minium de bon fonctionnement en application

Les temps minimum de bon fonctionnement ont été calculés pour des deux mémoires dans les conditions de l'application.

Le composant A possède un TMBF calculé pour 1% des échantillons défaillants, compris entre 10 et 21 ans. Ces valeurs sont en deçà des spécifications en durée de vie de l'application fixées elle à 30 ans. En revanche, le composant B possède un TMBF à 1 % qui est 2,6 fois supérieur à la durée de vie requise de l'application. Cette différence de TMBF vis-à-vis du mécanisme d'électromigration est expliquée par la différence de matériaux utilisés pour la réalisation des métallisations BEOL. De ce fait, il est possible de conclure d'après les rapports de qualification que la mémoire flash A est plus appropriée que la mémoire B dans pour l'application étudiée (Tableau 4.6).

Composant	TMBF _{APPLI_1%}	TMBF _{APPLI_63,2%}
Α	7 – 14 ans	21 – 45 ans
В	80 – 170 ans	250 – 550 ans

Tableau 4.6: Valeur des TMBF des deux mémoires flash 50 nm calculés en application pour 1% et 63,2 % d'échantillons défaillants et comparaison vis-à-vis de la spécification en durée de vie de l'application de 30 ans. En vert sont représentés les valeurs de TMBF supérieures à 30 ans et en rouge les valeur en dessous

Toutefois, les résultats ne signifient pas pour autant que le temps de défaillance de mémoire A soit comprise entre 7 et 14 ans. La méthodologie permet seulement d'indiquer que les 1 000 heures de test à 125°C pratiqués sur les 87 mémoires par le fabricant A ne suffisent pas à couvrir la durée de l'application. Pour être représentatif de la durée l'application étudiée, il aurait été nécessaire d'effectuer ce test sur des durées plus longues, à savoir 3 100 heures dans les conditions identiques ou alors 1 300 heures à 150°C.

Enfin, le cas d'étude présenté dans ce manuscrit s'est focalisé sur le mécanisme d'électromigration. Néanmoins, pour être complet, il est nécessaire d'appliquer la même méthodologie sur les autres mécanismes de défaillance jugés comme pertinents tels les porteurs chauds par exemple.

4.3.3 Synthèse

La méthodologie présentée permet d'évaluer un temps minimum de bon fonctionnement d'un composant dans une application spécifique, et cela vis à vis des mécanismes de dégradation prépondérants. Contrairement à la notion de temps de défaillance, le TMBF ne permet pas d'effectuer une prédiction de durée de vie réelle mais permet de calculer la durée moyenne de bon fonctionnement entre défaillances garantie par les tests du fabricant. De ce fait, elle permet d'évaluer la pertinence des tests de qualifications réalisés basée sur la probabilité d'apparition des défaillances. Dans le cas où ils ne le seraient pas, elle indique à un utilisateur de composant, les conditions opératoires à appliquer pour garantir le bon fonctionnement d'un composant.

Enfin, la méthodologie a été développée de façon à ce qu'elle puisse être utilisée dans l'industrie dans le domaine de la sélection composant. Elle permet à un utilisateur de choisir parmi plusieurs références commerciales, les composants les plus aptes à satisfaire les critères de fiabilité dans un environnement identifié et pendant une durée spécifiée.

4.4 Conclusions

Ce chapitre a permis de donner à un utilisateur des techniques et des outils d'évaluation de la fiabilité d'un composant fortement submicronique, non plus au niveau du silicium mais au niveau d'une référence commerciale standardisée. Dans ce but, deux approches ont été étudiées.

La première approche a été développée grâce à des tests de vieillissement accélérés. Ces tests permettent de reproduire sur une courte période, le comportement du composant qui pourrait être utilisé dans une application avionique ou aérospatiale réelle de longue durée. Dans cette optique, trois techniques de vieillissement de composants ont été explorées sur une mémoire SRAM de génération 90 nm. La première technique a consisté à prendre des contacts directement sur cellules mémoire à l'aide des nano pointes d'un probeur spécifique et d'extraire les caractéristiques de transistors NMOS de décharge. Cette technique s'est montrée fortement limitée car elle ne peut être utilisée pour réaliser des tests accélérés de vieillissement en raison de la faible durée de vie des pointes. De même, la seconde technique de tests que nous avons tentée a été basée sur l'utilisation des plots de contact FIB, où la grande variabilité introduite lors de la préparation des échantillons n'a pas permis d'obtenir des résultats cohérents et reproductibles.

En revanche, la troisième technique fondée sur l'association du testeur de composant ELECTRE et d'un système de contrainte thermique locale, a été beaucoup plus satisfaisante. Cette technique a permis de mettre en évidence une dégradation de type HCI sur une SRAM 90nm et d'en extraire la durée de vie associée. Les résultats montrent que pour une application aéronautique définie par le profil de mission n°1, le composant ne présente pas de risque de défaillance vis-à-vis de ce mécanisme. Enfin contrairement aux deux autres types d'expérimentations, cette technique est peu onéreuse et peut être appliquée de façon standard sur plusieurs familles de composants fortement submicroniques.

Les techniques de vieillissement correspondent à une approche de type recherche avec laquelle il est possible d'extraire des tendances quant à l'évolution de la durée de vie des composants. Toutefois, cette approche est assez lourde à mettre en place pour un utilisateur de composants qui ne disposerait pas de moyens de test. C'est pourquoi une méthodologie de sélection de composant exploitant les données de tests pratiqués et répertoriés chez le fabricant a été développée pour palier à cette difficulté sur deux technologies de mémoires Flash 50nm. Cette méthode permet de calculer le temps minimum de bon fonctionnement d'un composant fortement submicronique en se fondant sur les principaux éléments provenant de la technologie mis en adéquation au profil de mission requis.

A ce titre, un exemple concret d'application industrielle a permis de souligner la pertinence de cette approche étant donné la grande variété de technologie de composants disponibles dans le commerce et le grand nombre d'applications possibles.
CONCLUSION GENERALE ET PERSPECTIVES

Les composants commerciaux fortement submicroniques sont utilisés dans les équipements aéronautiques depuis quelques années pour des applications spécifiées pour durer plusieurs décennies. Néanmoins, ces composants sont destinés aux marchés de masse que représentent les secteurs de la micro-informatique et des télécommunications et non spécifiquement pour le marché aéronautique. De ce fait, ces composants ont spécialement été conçus pour des besoins dits de haute performance ou de basse consommation et ou la fiabilité n'est pas un critère prioritaire. Pour satisfaire ces marchés, une nouvelle génération de composants numériques émerge sur le marché tous les 2 ans, en imposant une diminution des dimensions des métallisations BEOL ainsi que des transistors FEOL et/ou l'introduction de nouveaux matériaux. Cependant, cette diminution est bien souvent non-homothétique, ce qui d'engendre une augmentation des contraintes en champ électrique et des courants de fuite au sein même des briques BEOL et FEOL du circuit intégré. Ces modifications ont conduit à une aggravation des mécanismes de défaillance par usure pour les générations de composants fortement submicroniques. D'autre part, l'approche classique qui consiste à s'assurer de la fiabilité d'un composant à l'aide des tests accélérés de qualification standard n'est désormais plus suffisante et donne des résultats parfois non valables. En effet, l'estimation de la durée de vie d'un composant nécessite de prendre en compte les mécanismes actifs de défaillance et d'usure, l'impact du niveau d'intégration de la technologie et finalement, de tenir compte du profil de mission réel de l'équipement dans lequel il est assemblé.

L'état de l'art a permis de recenser les mécanismes de défaillance d'usure pouvant potentiellement jouer un rôle important dans la durée de vie des composants en application fortement submicroniques. En ce sens, la physique de la défaillance des principaux mécanismes d'usure a été étudiée, à savoir : *l'Electromigration*, le *Stress Migration* et *le claquage d'oxyde intermétallique* pour l'empilement BEOL ainsi que *l'instabilité de la tension de seuil en température, les porteurs chauds* et *le claquage d'oxyde de grille* pour la partie FEOL. De plus, les lois d'endommagement associées a ces mécanismes ainsi que les paramètres des modèles correspondant aux technologies disponibles sur le marché des composants numériques en 2011, ont été proposés dans cette étude. Dans cette optique, cette étude a permis d'estimer la durée de vie moyenne d'un composant fortement submicronique (pour 63,2% des échantillons) quelle que soit la technologie et son utilisation. Néanmoins, la modélisation du mécanisme Stress Migration demande à être approfondie en considérant les contraintes de température et de cyclage. De même, les technologies silicium sur isolant ainsi que les matériaux III-V n'ont pas été investigués et ni modélisés.

Une estimation de la durée de vie moyenne des technologies DSM existantes a pu être déterminée au niveau silicium pour les nœuds 0,5 µm à 45 nm et pour trois profils d'application aéronautiques. Cette approche détaillée est fondée sur les mécanismes de dégradation préalablement modélisés. Notre étude fait apparaître plusieurs éléments importants de conclusions :

- Tout d'abord, outre le fait que la durée de vie des technologies numériques diminue avec la réduction des dimensions BEOL et FEOL, il apparait que l'accélération traduisant une nette dégradation de la fiabilité globale du composant reste très dépendante des matériaux utilisés en fonderie. L'étude a permis en ce sens d'identifier pour une application aéronautique fixée, la technologie et les matériaux du composant (via le fabricant) nécessaires pour respecter les spécifications en durée de vie.
- D'autre part, à technologie fixée, la tenue en durée de vie du composant dépend fortement de son type d'utilisation et du profil de mission qu'il lui est imposé. En ce sens, un facteur de sollicitation, noté α, correspondant au ratio de sollicitations des métallisations et des transistors sur le temps des équipements en fonctionnement a été défini. Par exemple, le mécanisme d'instabilité en température (BTI) a été identifié comme critique pour les structures CMOS, particulièrement pour les étages de charges composés des transistors PMOS et constitués d'oxydes de grille nitrurées. Ce point s'est montré en effet critique si ces dispositifs sont utilisés plus de 10% du temps et que le calculateur est en fonctionnement dans des conditions thermiques définies par le profil n°1. Ceci n'est pas le cas pour les deux autres profils de missions étudiés.
- Enfin, dans le cas où un composant ne satisferait pas ses spécifications en durée de vie, il est possible d'adapter les conditions d'utilisations telles que la fréquence de sollicitation α ou encore les conditions thermiques, de façon à fiabiliser un produit qui ne le serait pas sans cela. Ce type de solution peut se traduire par exemple par une modification de l'architecture du système telle que la mise en place de redondance

(matérielle ou logicielle), de triplication ou encore de solutions de refroidissement thermique adaptées.

Cette étude a permis de fournir aux utilisateurs de composants numériques CMOS de génération 0,5 µm à 45 nm, plusieurs leviers pour qu'ils puissent s'assurer d'une marge de sécurité à travers la détermination de la durée de vie de leur produit, à savoir : la technologie des composants comprenant le nœud et les matériaux, le fabricant ainsi que les conditions d'utilisations. Aussi, étant donné le nombre important de leviers, un outil informatique a été développé au cours de ces travaux pour répondre aux besoins des utilisateurs EADS. Une amélioration potentielle de cette approche consisterait à prendre en compte les incertitudes et la variabilité des paramètres d'entrées des modèles comme par exemple les énergies d'activations des mécanismes en température ou encore les dimensions des structures, les médianes et écarttypes de résultats de stress suivants les différents types de contraintes. Ceci permettrait d'enrichir le calcul analytique en sortie en calculant une densité de probabilité du temps de défaillance incluant des effets statistiques. Un autre axe de développement à envisager serait d'inclure les nouvelles technologies en dessous des nœuds 40 nm telles que les structures de types Silicon On Insulator "Fully Depleted" (FDSOI) ou "Partially Depleted" (PDSOI) qui sont associées aux modifications majeures représentées par l'introduction de diélectriques de grille à haute constantes (Hi-K), de grille métallique, de couches implantées de SiGe, tout comme les structures à substrats étirés, à trois dimensions (3D) de type FinFET et l'utilisation de nouveaux matériaux III-V pour toutes ces nouvelles architectures émergentes.

L'outil de prédiction développé chez EADS permet d'estimer la durée de vie d'un composant en fonction des technologies et du profil de mission. Cependant, nous devons souligner que cette approche a été très orientée au niveau silicium (wafer level) et qu'elle doit être validée par plusieurs méthodes complémentaires. Pour cela, trois voies ont été potentiellement investiguées, à savoir la voie expérimentale, des simulations et la voie technologique.

Parmi les techniques explorées, les tests par vieillissement de composants commerciaux à l'aide d'un système électrique carte mère/carte fille et d'un système de contrainte thermique locale, semble la technique la plus standard pour un utilisateur de composant. De plus, elle demeure la plus reproductible et la moins coûteuse des solutions envisagées dans ce manuscrit,. Ces tests ont notamment permis de mettre en évidence une défaillance par porteurs chauds sur une SRAM de génération 90 nm, bien que celle-ci ait passé avec succès les tests de qualification. Ceci souligne le fait que les tests de qualifications standards pratiqués chez les fabricants au

niveau du produit, ne permettent pas d'évaluer la criticité des mécanismes d'usure du circuit intégré pour des profils de missions avioniques.

De plus, ces tests ne permettent pas non plus d'évaluer la durée de vie d'un dispositif en application car cela supposerait que les fabricants connaissent le profil de mission de l'utilisateur, ce qui n'est en principe pas le cas. Il s'offre donc deux possibilités :

- Une meilleure coopération et communication entre les fabricants de composants et les utilisateurs finaux. Les domaines d'échanges souhaités sont principalement les données technologiques, les résultats de tests pratiqués en fonderie, les paramètres et les lois de dégradation utilisées et enfin des informations relatives aux effets de la variabilité des procédés de fabrication sur les paramètres représentatifs des performances et qui subiront les conséquences des phases de stress.
- Dans le cas où une collaboration serait difficilement envisageable, il devient nécessaire de pratiquer des tests de vieillissement chez l'utilisateur. Ceci implique pour une technologie donnée, la réalisation de tests dans des conditions thermiques et électriques différentes afin de pouvoir extraire un facteur d'accélération. Cette approche est réalisable à l'aide d'un testeur comme celui utilisé aux cours de ces travaux. Néanmoins, elle nécessite des temps conséquents d'expérimentations et d'exploitations qui ne sont pas forcément accessibles à un utilisateur qui ne disposerait pas de tels moyens de tests.

L'approche expérimentale qui a été décrite dans ce manuscrit permet d'une part, de valider les modèles de prédictions utilisés et ajustés au niveau du silicium et d'autre part, d'estimer un temps de défaillance moyen d'un composant fonctionnel en application. Toutefois, cette approche montre ses limites si les conditions expérimentales d'accélération ne sont pas assez sévères pour provoquer une défaillance dans un temps raisonnable. C'est pourquoi, nous avons développé de façon complémentaire, une approche méthodologique qui permet de calculer le temps minium de bon fonctionnement d'un composant spécifique à une application (TMBF) garanti au travers des résultats de tests de qualification du fabricant, en fonction de l'utilisation et de la technologie. Cette notion de TMBF vient compléter la notion de temps de défaillance (ITFF) dans le sens où il n'est plus nécessaire de connaitre le temps effectif de "fin de vie" du

composant pour savoir s'il remplira toutes ses spécifications de fiabilité pour ses missions réelles, c'est à dire sous fortes contraintes aéronautique et aérospatiales.

Cette approche a été appliquée au cours des trois années de cette thèse sur plusieurs familles de composants et pour diverses applications aéronautiques typiques. Une autre perspective intéressante qui découle des remarques précédentes serait de coupler le calcul du TMBF qui traduit la notion d'usure des composants fortement submicroniques, à l'estimation du taux de défaillance en application décrit dans le guide FIDES **[FID 09]**. Ceci pourrait concourir à l'établissement de nouvelles normes de fonctionnements et d'utilisations, pour un client potentiel qui chercherait à faire évoluer des applications spécifiques en toute sécurité vers des technologies intégrées.

PUBLICATIONS ET ACTIVITES

Revues à comité de lecture :

A New Policy for COTS Selection: Overcome the DSM Reliability Challenge F. Molière, B. Foucher, P. Perdu, A. Bravaix, *SAE International Journal of Aerospace*, vol 4, issue 2, p 1475-1484, November 2011

Reliability Issues of Electronics in Aeronautics B. Foucher, F. Molière Ouvrage *Smart Revue Integration and Reliability*, p 402-410, ISBN 978-3-932434-77-8, 2010

Comment déterminer la fiabilité d'un composant fortement sub micronique (DSM) Partie 1 : Impact de l'intégration et des choix technologiques sur la fiabilité d'un composant DSM F. Molière CCT CNES numéro 33, 1^{er} Septembre 2010

Use of Laser to Explain Heavy Ion Induced SEFIs in SDRAMs A.Bougerol, F. Miller, N. Guibbaud, R. Gaillard, F. Moliere, and N. Buard *IEEE Transactions on Nuclear Science*, vol. 57, no. 1, february 2010

Analysis of deep submicron VLSI technological risks: A new qualification process for professional electronics

F. Molière, B. Foucher, P. Perdu, A. Bravaix, *Microelectronics Reliability*, vol 49, p1381–1385, 2009

Conférences, symposiums à comité de lecture :

A New Policy for COTS Selection: Overcome the DSM Reliability Challenge F.Molière, B.Foucher, P.Perdu, A.Bravaix *Proceedings of SAE International* Présenté à la conférence SAE le 21 Octobre 2011

Présentations en Conférence et Workshop :

The issus of Long Term Reliability for Industrial application: Deep sub micron Challenges and perspectives, F. Molière *WorkShop LOTUS*, 9th October 2009, Arcachon

Poster:

Analysis of deep submicron VLSI technological risks: A new qualification process for professional electronics

F. Molière, B. Foucher, P. Perdu, A. Bravaix, *ESREF 2009*, 7th October 2009, Arcachon

Brevets internationaux:

Method for estimating the lifespan of a deep-sub-micron integrated electronic circuit Inventeurs: F.Molière, B.Foucher Publication n°: WO/2011/000888, FR/2009/0054496

Procédés et dispositifs de mise sous contrainte d'un circuit intégré

Inventeurs: F.Molière, S.Morand, A.Douin, D.Peyre, C. Binois, G.Salvaterra (Date du dépôt : 20/04/2010, Numéro de dépôt : 10 52978)

REFERENCES BIBLIOGRAPHIQUES

- [AIR 10] Airbus SAS, présentation interne groupe EADS, communication privé.
- [ALA 05] M. A. Alam, S. Mahapatra, "A comprehensive model for PMOS NBTI degradation", Microelectronics Reliability 45, p 71-81, 2005.
- [ALA 07] M. A. Alam, H. Kufluoglu, D. Varghese, S. Mahapatra, "A comprehensive model for PMOS NBTI degradation: recent progress", Microelectronics Reliability 47, p 853-863, 2007.
- [AON 05] H. Aono, E. Murakami, K. Okuyama, A. Nishida, M. Minami, "Modeling of NBTI saturation effect and its impact on electric field dependence of the lifetime", Microelectronics Reliability 45, p 1109 – 1114, 2005.
- [ARI 05] R. Arijit, C-M. Tan, R. Kumar, X-T Chen, "Effect of test condition and stress free temperature on the electromigration faillure of Cu dual damascene submicron interconnect line-via test structures", Microelectronics Reliability 45, p1443-1448, 2005.
- [BAO 03] L. Baozhen, T. Sullivan, T. Lee, D. Badami, "Reliability challenges for copper interconnect", Microelectronics Reliability 44, p365-380, 2004.
- [BEI 98] I. Beinglass, M. Naik, "CVD Al/PVD Al integration for advanced via and interconnect technology", Thin Solid Films, p35-44, 1998.
- [BER 00] L.K Bera, B. Senapati, S. Maikap, C.K. Maiti, "Effects of O2/N2O-plasma treatment on strained Si, Solid State Electronics 44", p1533-1536, 2000.
- [BIE 04] L. Biesemans, K. Schepers, K. Vanstreels, J. D'Haen, W. De Ceuninck and M. D'Olieslaeger, "MTF test system with AC based dynamic joule correction for electromigration tests on interconnects", Microelectronics Reliability 44, p1849-1854, 2004.
- [BLA 67] J. R. Black, "Mass Transport of Aluminum by Momentum Exchange with Conducting Electrons", 6th Reliability Physics Symposium, 1967.
- [BLE 76] A. Blech, "Electromigration in thin aluminum films on titanium nitride", J. Appl. Phys. 47 1203, 1976.
- [BOH 07] M-T. Bohr, R-S. Shau, K. Mistry, "The High K solution", IEEE Spectrum p24-29, novembre 2007.
- [BOU 97] A. Bouhdada, S. Bakkali, A. Touhami, "Modelling of Gate-Induced drain leakage in relation to technological parameters and temperature", Microelectronic Reliability 37, p 649-652, 1997.
- [BRA 03a] A. Bravaix, "Modélisation de MOSFETS Spice et BSIM", Cours ISEN, 2003.
- [BRA 03b] A. Bravaix, C. Trapes, D. Goguenheim, N. Revil, E. Vincent, "Carrier injection efficiency for the reliability study of 3.5-1.2nm thick gate-oxide CMOS technologies", Microelectronic Reliability, vol. 43, p1241-1246, 2003.
- [BRA 09] A. Bravaix, C. Guerin, V. Huard, D. Roy, J.-M. Roux., E. Vincent, "Hot-Carrier Acceleration Factors for Low Power Management in DC-AC stressed 40 nm NMOS node at High Temperature", IEEE International Reliability Physics Symposium (IRPS) Proc., p 531-548, 2009.
- [CAR 03] E. Cartier, L. Pantisano, A. Kerber, G. Groeseneken, "Correlation between charge injection and trapping in SiO₂/HfO₂ gate stacks", presented at IEEE INFOS, 2003.

- [CHA 04] S. Chakravarthi, A.T. Krishnan, V. Reddy, C.F. Machala and S. Krishnan, "A Comprehensive Framework For Predictive Modeling of Negative Bias Temperature Instability", IEEE 04CH37533 4P Annual International Reliability Physics Symposium, Phoenix, 2004.
- [CHA 05a] R. Chau, J. Brask, S. Datta, G. Dewey, M. Doczy, B. Doyle, J. Kavalieros, B. Jin, M. Metz, A. Majumdar, M. Radosavljevic, "Application of High-k Gate Dielectrics and Metal Gate Electrodes to enable Silicon and Non-Silicon Logic Technology", Microelectronic Engineering, Vol 80, Iss 1, 2005.
- [CHA 05b] P. Chaparala and D. Brisbin, "Impact of NBTI and HCI on PMOSFET threshold voltage drift", Microelectronics Reliability 45, p 13-18, 2005.
- [CHE 02] G. Chen, M.F. Li, H. Ang, J.Z.Zheng and D.L. Kwong, "Dynamic NBTI of P-MOS Transistors and its impact on MOSFET Scaling", IEEE Electron device letters, vol 23, n°12, december 2002.
- [CHE 04] K. H.Cheng and Ahila Krishnamoorthy, "Effect of ramp rate on dielectric in CU-SiOC interconnects", Thin Solid Films 462-463, p 316-200, 2004.
- [CHE 07] F. Chen, P. McLaughlin, J. Gambino, E. Wu, J. Demarest, D. Meatyard and M. Shinosky, "The effect of Metal Area and Line Spacing on TDDB Characteristics of 45 nm low k SiCOH Dielectrics", Solid State Electronics 51, p 565-571, 2007.
- [CHO 08] N.A. Chowdhury, G. Bersuker, C. Young, R. Choi, S. Krishanan, D. Misra, "Breakdown characteristics of nFETs in inversion with metal/HfO2 gate stacks", Microelectronics Engineering 85, p 27-35, 2008.
- [CHU 88] J.E. Chung, M.-C. Jeng, G. May, P.K. Ko, and C. Hu, "Hot-electron currents in deep submicrometer mosfets", IEDM Tech, 1988.
- [CHU 90] J.E. Chung, M.-C. Jeng, J.E. Moon, P.K. Ko, C. Hu, "Low-Voltage Hot-Electron Currents and Degradation in Deep-Submicrometer MOSFET's", IEEE Trans. Elec. Dev., Vol. 37, pp. 1651-1657, 1990.
- [CIM 09] Document interne CIMPACA, 2009.
- [DEA 67] Deal B. E., Sklar M., Grove A.S., Snow E. H., "Characterization of the Surface-State Charge (Qss)of the thermally Oxidized Silicon", Journal of electrochemical society, solid state science, Volume 114, Number 3, pp. 266-274, 1967.
- [DEG 05] R. Degraeve, B. Govoreanu, B. Kaczer, J. Van Houdt, G. Groeseneken, "Measurement ans statistical analysis of single trap current-voltage characteristics in ultra-thin SiON", IEEE 05CH37616 43th Annual International Reliability Physics Symposium, San Jose, 2005.
- [DEN 04a] M. Denais, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, N. Revil, A. Bravaix, "Interface Trap generation and hole trapping under NBTI and PBTI in advanced CMOS Technology with a 2nm gateoxide", Transactions on Device and Materials Reliability, Vol. 4, pp. 715-722, 2004.
- [DEN 04b] M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, Y. Rey-Tauriac, N. Revil, "Onthe-fly" characterization of NBTI in ultra-thin gate-oxide PMOSFET's", IEEE International Electron Device Meeting (IEDM) Proc., pp. 109-112, 2004.
- [DEO 08] S. Deora and S. Mahapatra, "A Study of NBTI in HfSiON/TiN p-MOSFETs Using Ultra-Fast On-The-Fly (UF-OTF) IDLIN Technique", Physical and Failure Analysis of Integrated Circuits, 2008.
- [FID 09] Guide FIDES, "Méthodologie de fiabilité pour les systèmes électroniques", Mai 2009.
- [FOR 09] L. Forli, CIMPACA, communication privé, 2009.
- [GAR 04] X. Garros, C. Leroux, G. Reimbold, J. Mitard, B. Guillaumot, F. Martin, J-L Autran, "Reliability assessment of ultra-thin HfO2 oxides with TiN gate and polysilicon-n+ gate", IEEE International Reliability Physics Symposium Proceedings, N°42nd, pp. 176 180, 2004.

[GAR 09]	X. Garros, M. Casse, G. Reimbold, M. Rafik, F. Martin, F. Andrieu, V. Cosnier, F. Boulange, "Performance and reliability of advanced High-K/Metal gate stacks", Microelectronic Engineering 86, p1609–1614, 2009.
[GHA 03]	T. Ghani, et al, "A 90nm High Volume Manufacturing Logic Technology Featuring Novel 45 nm Gate Length Strained Silicon CMOS Transistors", International Electron Devices Meeting, December 9, 2003.
[GIG 03]	L.M. Gignac, CK. Hu, E.G. Liniger, "Correlation of electromigration lifetime distribution to faillure mode in dual Damascene Cu/SiLK interconnects", Microelectronic Engineering 70, p398-405, 2003.
[GUE 08]	C. Guérin, "Etude de la dégradation par Porteurs Chauds des Technologies CMOS Avancées en fonctionnement Statique et Dynamique", Thèse de doctorat de 3 ^{ième} cycle de l'Université d'Aix Marseille I, soutenue le 10 Octobre 2008.
[GUE 09]	C. Guérin, V. Huard, A. Bravaix "General framework about defects creation at the Si/SiO ₂ interface", Journal of Applied Physics, Vol. 105, pp. 114513-114524, 2009.
[GUI 05]	J-F. Guillaumond, "Etude de la résestivité et de l'électromigration dans les interconnexions déstinées aux technologies des nœuds 90 à 32 nm", Université Joseph Fourier Grenoble I, thèse de doctorat soutenue le 2 décembre 2005.
[GUT 95]	R-J. Gutmann, T-P. Chow, A-E. Kaloyeros, W-A. Landford, S-P. Muraka, "Thermal stability of on chip copper interconnect structures", Thin Solid Films 262, p177-186, 1995.
[HE 05]	Y. He, M. Xu, C. Tan, "Effects of plasma nitridation on ultra-thin gate oxide electrical and reliability characteristics", Solid-State Electronics 49, p57-61, 2005.
[HU 03]	C.K. Hu, L. Gignac, R. Rosenberg, E. Liniger, J. Rubino, C. Sambucetti, A. Stamper, A. Domenicucci, X. Chen, "Deduced Cu Interface diffusion by CoWP surface coating, Microelectronic Enginneering 70, p 406-41, 2003.
[HU 06]	C-K. Hu, L. Gignac, R. Rosenberg, "Electromigraton of Cu/low dielectric constant interconnects", Microelectronics Reliability 46, p213-231, 2006.
[HU 85]	C. Hu S. Te, F-C Hsu, M P-K Ko, T-Y Chan, W. Terrill, "Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement", IEEE transactions on electron devices, vol. ed-32, no. 2, february 1985.
[HUA 05]	V. Huard, Denais M., Perrier F., Revil N., Parthasarathy C., Bravaix A., Vincent E., "A thorough investigation of MOSFETs NBTI degradation", Microelectronic Reliability, Vol. 45, N°1, pp. 83-98, 2005.
[HUA 07]	V. Huard, C. Parthasarathy, N. Rallet, C. Guerin, M. Mammasel, D. Barge, C. Ouvrard, "New characterization and modeling approach for NBTI degradation from transistor to product level", Electron Devices Meeting, IEDM 2007. IEEE International, 2007.
[HUA 11]	V. Huard, "From device to library reliability in advanced CMOS nodes", Tutorial IRPS, Monterey, 2011
[ITR 07]	"International Technology Roadmaps for Semiconductor", edition 2000-2007.
[JEO 09]	Y-W Jeon, D-H Ka, C-G Yu, W-J Cho, M-S Islam, J-T Park, "NBTI and hot carrier effect of SOI p-MOSFETs fabricated in strained Si SOI wafer", Microelectronics Reliability 49, p994–997, 2009.
[JEP 11]	JEDEC JEP122E, 2011.
[JEP 77]	K.O. Jeppson, C.M. Svensson, "Negative bias stress of MOS devices a high electric fields and degradation of MNOS devices", J. Appl. Phy. Vol. 48, pp. 2004-2014, 1977.
[JIA 05]	L. Jiang, "Hafnium-Doped Tantalum Oxide High-k Gate Dielectric Films for Future CMOS Technology", Université du Texas A&M, Thèse de doctorat soutenue en décembre 2005.

[KAR 99]	S.P. Karna, A.C. Pineda, R.D. Pugh, W.M. Shedd, T.R. Oldham, "Electronic Structure and Theory and Mechanisms of the Oxide Trapped Hole Annealing Process", IEEE Trans. Nuclear Sc., Vol. 47, pp. 2316-2324, 1999.
[KAU 02]	T. Kauerauf, R. Degreave, E. Cartier, B. Govoreanu, P. Blomme, B. Kaczer, L. Pantisano, A. Keber, G. Groeseneken, "Towards Understanding degradation and breakdown of SiO ₂ /High-k stacks", IEEE International Electron Device Meeting, p 521-524, 2002.
[KER 03]	A. Kerber., E. Cartier, L. Pantisano, T. Kaureauf, Y. Kim, A. Hou, G. Groeseneken, U. Schwalke, "Origin of the threshold voltage instability in SiO ₂ /HfO ₂ Dual layer gate dielectrics", IEEE Electron Device Letters, Volume 24, Number 2, pp. 87-89, 2003.
[KIM 07]	J. Kim, E-T. Ogawa and J-W. McPherson, "Time Dependent Dielectric Breakdown Characteristics of Low-k Dielectric (SiOC) Over a Wide Range of Test Areas and Electric Fields", IEEE 07CH37867 45th International Reliability Physics Symposium, Phoenix, 2007.
[KNO 07]	S. Knowles, "The return of silicon efficiency", 18th IEEE Symposium on computer arithmetic, 2007.
[KOI 04]	N. Koike and K. Tatsuuma, "A Drain Avalanche Hot Carrier Lifetime Model for n- and p-Channel MOSFETs", IEEE Transactions On Device And Materials Reliability, Vol. 4, no. 3, pp. 457-466, September 2004.
[KOY 86]	M. Koyanagi, A.G. Lewis, J. Zhu, R.A. Martin, T.Y. Huang, J.Y. Chen, "Investigation and Reduction of Hot-Electron Induced Punchthrough (HEIP) Effects in Submicrometer p-MOSFETs", IEDM Tech. Dig., p. 722-725, 1986.
[LEE 08]	Y-M. Lee, Y. Wu, "Influence of Si/SiO2 interface properties on electrical performance and breakdown characteristics of ultrathin oxide/nitride dielectric films", Applied Surface Science 254, p4591-4598, 2008.
[LIN 07]	M.H. Lin, K.P. Cang, KC. Su, Tahui Wang, "Effects of width scaling and layout variation on dual damascene copper interconnect electromigration", Microelectronics Reliability 47, p2100-2108, 2007.
[LLO 04]	J.R Lloyd, M.R. Lane, XH. Liu, E.Liniger, T.M. Shaw, CK. Hu, R. Rosenberg, Realiability Challenges with Ultra-low K Interlevel Dielectrics, Microelectronics Reliability 44, pp1835-1841, 2004
[LLO 07]	J.R. Lloyd, "Black'sLaw revisited-Nucleation and growth in electromigration failure", Microelectronics Reliability 47, p 1468-1472, 2007.
[LOP 04]	L. Lopez, P. Masson, D. Nee, R. Bouchakour, "Temperature and drain voltage dependence of gate- induced drain leakage", Microelectronic Engineering 72, p101-105, 2004.
[MAN 98]	Y. Maneglia, "Analyse en profondeur des défauts de l'interfaceSi-SiO2 par la technique du pompage de charge", Université Polytechnique de Grenoble, Thèse de doctorat soutenue le 18 Décembre 1998.
[MAR 07]	A. Marras, M. Impronta, I. De Munari, M.G. Valentini and A. Scorzoni, "Reliability assessment of multi-via Cu-damascene structures by water-level isothermal electromigration tests", Microelectronics Reliability 47, p1492-1496, 2007.
[MAS 07]	P. Masson, "Panorama des mémoires innovantes sur silicium", Cours Master MINELEC, Source Laboratoire In2MP, 2007.
[MAT 10]	S. Matsumoto and Al, "Highly manufacturable ELK integration technology with metal hard mask process for high performance 32 nm-node interconnect and beyond", Interconnect Technology Conference (IITC), Burlingame CA, June 2010.
[MCP 02]	J-W. McPherson, A. Shanware and L. Colombo, "Electrical and Reliability Characteristics of HfSiON Gate Dielectric", Topical Research Conference on reliability, Austin, Texas, October 28-30, 2002.
[MCP 03]	J-W. McPherson, J. Kim, A. Shanware, H. Mogul and J. Rodriguez, "Trends in the Ultimate Breakdown Strength of High Dielectric Constant Materials", IEEE Transaction on electron devices vol 50, n°8, august 2003.

[MCP 87]	J-W. McPherson and C-F Dunn, "A model for stress: induced metal notching and voiding in very large: scale: integrated Al–Si (1%) metallization", Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures, vol 5 issue 5, p1321-1325, 1987.
[MCP 98]	J-W. McPherson, "Disturbed bonding states in SiO/sub 2/ thin-films and their impact on time-dependent dielectric breakdown", IRPS 1998, pp. 47-56, 1998.
[MIS 07]	K. Mistry et al., "A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193 nm Dry Patterning, and 100% Pb-free Packaging," IEDM Tech. Dig., pp. 247-250, 2007.
[NIC 02]	C-J. Nicklaw, Z-Y. Lu, D-M. Fleetwood, R-D. Schrimpf, S-T. Pantelides, "The Structure, Properties, and Dynamics of Oxygen Vacancies in Amorphous SiO ₂ ", IEEE Trans. On Nuclear Sc., Vol. 49, N° 6, pp. 2667-2673, 2002.
[NTR 97]	The National Technology Roadmap for Semiconductors, Technology Needs, Semiconductor Industry Association, edition 1997.
[OGA 01]	E-T. Ogawa, A-J. Bierwag, K-D. Lee, H. Matsuhashi, P-R. Justison, A-N. Ramamurthi, and P-S. Ho, "Direct observation of a critical length effect in dual-damascene Cu/oxide interconnects", Appl. Phys. Lett. 78, 2652, 2001.
[OGA 02]	E-T. Ogawa, J-W. McPherson, J-A. Rosal, K-J. Dickerson, T-C. Chiu, L-Y. Tsung, M-K. Jain, T-D. Bonifeld, J-C. Ondrusek and W-R. Mckee, "Stress-Induced Voiding Under Vias Connected To Wide Cu Metal Leads", IEEE 02CH37320, 40 th Annual International Reliability Physics Symposium, Dallas, Texas, 2002.
[PEN 01]	C. Pennetta, L. Reggiani, Gy. Trefan, F. Fantini, A. Scorzoni, I. De Munari, "Investigation of the role of compositional effects on electromigration damage of metallic interconnects", Computational Materials Science 22, p13-18, 2001.
[PER 06]	P. Perdu, "Long term reliability issues in the ddep-submicron range", Aging and Self Life of NANO/MEMS Technology Workshop, 2006.
[PET 06]	C. Petit and D. Zander, "Low voltage induced leakage current and time to breakdown in ultra thin oxides (1,2-2,3 nm) ", Microelectronics Reliability 47, p401-408, 2006.
[PIC 06]	D. Pic, D. Goguenheim, J-L. Ogier, "Long range statistical lifetime prediction of ultra-thin SiO2 oxides: inluence of accelerated ageing methods and extrapolation models", Microelectronics, 2006 25th International Conference on Volume, Issue, 14-17, p528 - 531, 2006.
[POB 10]	G. Pobegen, T.Aichinger, M.Nelhiebel and T. Grasser, "Dependence of the Negative Bias Temperature Instability on the Gate Oxide Thickness", IRPS, p 1073, 2010.
[POM 00]	T. Pompl, H. Wurzer, M. Kerber, I. Eisele, "Influence of gate oxide breakdown on MOSFET device operation", Microelectronics Reliability 40, p37-47, 2000.
[POR 06]	M. Porti, L. Aguilera, X. Blasco, M. Nafria and X. Aymerich, "Reliability of SiO2 and high-K gate insulator: A nanoscale study with conductive atomic force microscopy", Nanoscale Imaging and Metrology of devices and innovative materials, Proceedings of the European Material Reasearch Society 2006 syposium F, p 501-505, 2006.
[PRE 98]	J. Prendergast, N. Foley, J-S. Suehle, "Investigation of the Intrinsic Si02 Area Dependence Using TDDB Testing and Model Integration Into the Design Process", Microelectronics Reliability 38, p1121-1125, 1998.
[RAN 06]	R. Singh, "Reliability and performance limitations in SiC power devices", Microelectronics Reliability 46, p713-730, 2006.
[RIB 05]	G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, C. Leroux, G. Ghibaudo, "Review on High-K dielectrics reliability issues", Transactions on Device and Materials Reliability, Vol. 5, pp.5-19, 2005.

- [RIB 07] G. Ribes, M. Rafik and D. Roy, "Reliability isues for nano-scale CMOS dielectrics", Microelectronic Engineering 84, p 1910-1916, 2007.
- [SHA 99] G-G. Shahidi, A. Ajmera, F. Assaderaghi, R-J. Bolam, E. Leobandung, W. Rausch, D. Sankus, D. Schepis, L-F. Wagner, K-Wu, and B. Davari, "Partially Depleted SOI Technology for Digital Logic", ISSCC Tech. Digest, p. 426-428, 1999.
- [STA 06] J.H. Stathis, S. Zafar, "The negative bias temperature instability in MOS devices: A review", Microelectronics Reliability, Volume 46, Issues 2-4, , p 270-286, February-April 2006.
- [SUN 01] J. Suné, "New physics-based analytic approach to the thin-oxide breakdown statistics", IEEE Electron. Dev. Lett., Vol. 22, no 6, p.296-298, 2001.
- [TAK 82] I. Takashi, N. Tetsuo, I. Hajime, "Advantage of thermal Nitride and Nitroxide Gate Films in VLSI Process", IEEE transactions on electron devices, vol. ed-29, no. 4, 1982.
- [TAK 83]E. Takeda, N. Suzuki, "An Empirical Model for Device degradation Due to Hot-Carrier Injection",
IEEE Elec. Dev. Lett. 4, pp. 111-114, 1983.
- [TAM 84]S. Tam, P.-K. Ko, C. Hu, "Lucky-Electron Model of Channel Hot-Electron Injections in MOSFET's",
IEEE Trans. Elec. Dev. 31, p. 1116, 1984.
- [TAU 98] Y. Taur, C.H. Wann, D. J. Frank, "25nm CMOS Design Consideration", IEDM Proc., pp. 789-792, 1998.
- [TOR 95] J. Torres, "Advanced copper interconnections for silicon CMOS technologies", Applied Surface Science91, p112-123, 1995.
- [TSU 05] S. Tsujikawa, J. Yugami, "Positive charge generation due to species of hydrogen during NBTI phenomenon in pMOSFETs with ultra-thin SiON gate dielectrics", Microelectronics Reliability 45, p 65–69, 2005.
- [TZO 86] J-J. Tzou, C-C. Yao, H-W-K. Chan, "Hot-Carrier-Induced degradation in P-Channel LDD MOSFETs", IEEE Electron Device Lett., Vol. 7, N°1, pp. 5-8, 1986.
- [VIN 07] E. Vincent, "Wafer Level Reliability", Cours Master MINELEC, Sources STMicroelectronics Crolles, 2007.
- [VIN 96] E. Vincent et al, "Electric field dependence of TDDB activation energy in ultrathin oxides", Microelectronics Reliability Vol. 36, p.1643, 1996.
- [VIT 06] J. Vitiello, "Etude des matériaux diélectriques à très faible permitivité déposé en phase vapeur développés pour l'isolation des interconnexions cuivre des circuits intégrés pour les générations technologiques 45 nm et 32 nm", INSA de Lyon, Thése de doctorat soutenue en 2006.
- [VOL 04] R-P. Vollertsen, E. Y. Wu, "Voltage acceleration and t63.2 of 1.6–10 nm gate oxides", Microelectronics Reliability, Volume 44, Issue 6, p 909-916, June 2004.
- [WON 07] H-Y. Wong, N-F. Mohd Shukor, N. Amin, "Prospective development in diffusion barrier layers for copper metallization in LSI", IEEE 07CH37867, 45th International Reliability Physics Symposium, Phoenix, 2007.
- [WU 03] C.T. Wu, A. Mieckowski, R.S Ridley Sr, G Dolny, T. Grebs, J. Linn, J. Ruzyllo, "Effect of nitridation on the reliability of thick gate oxides", Microelectronics Reliability 43, p43-47, 2003.
- [WU 05] E.Y. Wu and J. Suné, "Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability", Microelectronics and Reliability 45, p 1809-1834, 2005.
- [WU 08] W.Z. Yu, Y.Y. Tang, C.C. Chun, L. YueJin, W. JiaYou, L. Jing, L. Bin, "Temperature-dependent stressinduced voiding in dual-damascene Cu interconnects", Microelectronics Reliability 48, p578–583, 2008.

[XU 98]	J.P Xu, P.T. LAF, L. Huang, H.B. Lo and Y.C. Cheng, "Greatly suppressed stress-induced shift of GIDL in N2O based n MOSFET", Solid State Electronics 42, p1665-1669, 1998.
[YOU 06]	C-D. Young, R. Choi, B-H Lee, "Electron Trap Generation in High-K Gate Stacks by Constant Voltage Stress", EEE Transactions on device and materials reliability, vol 6, n°2, June 2006.
[YU 08]	W.Z Yu, Y. YinTang, C. ChangChun, L. YueJin, W. JiaYou, L. Jing, L. Bin, "Temperature-dependent stress-induces voiding in dual-damascene Cu interconnects", Microelectronics Reliability 48, p 578-583, 2008.
[YUA 08]	X. Yuan, J-E. Park, J. Wang, E. Zhao, D. C. Ahlgren, T. Hook, J. Yuan, V. W. C. Chan, H. Shang, C- H. Liang, R. Lindsay, S. Park, H. Choo, "Gate-Induced-Leakage current in 45nm CMOS Technology", IEEE Trans. on Device Material reliability, Vol. 8, No. 3, pp. 501-508, 2008
[ZHA 01]	J-H Zhao, W-J Qi, P Ho, "Thermomechanical property of diffusion barrier layer and its effect on the stress characteristics of copper submicron interconnect structures", Microelectronics Reliability, Volume 42, Issue 1, p27-34, 2002.
[ZHA 06]	L. Zhang, "Stress Induced Voiding in Dual-Damascene Cu interconnects", Term paper for EM 397: Thin Film Mechanics, 2006.
[ZHA 09]	J.F. Zhang, "Defects and instabilities in Hf-dielectric/SiON stacks (Invited Paper) ", Microelectronic Engineering 86, p 1883–1887, 2009.

LIENS INTERNET

[AMD 11]	http://www.xbitlabs.com/articles/cpu/display/athlon64-venice_2.html, consulté en Novembre 2011
[AVI 11]	Site internet http://www.aviation-fr.info/, consulté le 05/01/2011.
[HOF 10]	T.Y. Hoffmann, http://www.electroiq.com/articles/sst/2010/03/integrating-high-k.html, Mars 2010, consulté en Novembre 2010.
[IBM 07]	IBM, "High-K, Advanced material for futur CMOS generation", http://www.zurich.ibm.com, 2007.
[INT 03]	Intel, "Intel's High-k/Metal Gate Announcement", http://www.intel.com/, 2003.
[JED 11]	http://www.jedec.org/standards-documents, 2011.
[PTM 06a]	Predictive technology model, Beta Version of PTM for HCI, http://ptm.asu.edu/, document de 2006, consulté en 2010
[PTM 06b]	Predictive technology model, Beta Version of PTM for NBTI, http://ptm.asu.edu/, document de 2006, consulté en 2010
[VIR 02]	VirginSemi, "The General Properties of Si, Ge, SiGe, SiO ₂ and Si ₃ N ₄ ", http://www.virginiasemi.com, document de 2002, consulté en Mai 2010.

Fiabilité des technologies CMOS fortement submicroniques (DSM) pour des applications avioniques, aérospatiales et militaires

Résumé

Depuis ces dernières années, les composants fortement submicroniques du commerce sont utilisés dans les équipements aéronautiques pour des applications spécifiées pour durer plusieurs décennies. Toutefois, ces composants sont destinés aux marchés de masse que représentent les secteurs de la microinformatique et des télécommunications et ne sont pas spécifiques au marché aéronautique. De ce fait, ces composants sont spécialement conçus pour des besoins dits de haute performance ou de basse consommation et pour lesquels la fiabilité n'est pas un critère prioritaire. Pour satisfaire ces marchés, une nouvelle génération technologique émerge tous les deux ans en imposant à chaque fois, une diminution des dimensions des métallisations BEOL ainsi que des transistors FEOL et/ou l'introduction de nouveaux matériaux. Ces modifications ont conduit à une aggravation des mécanismes de défaillance par usure pour les générations de composants fortement submicroniques, au point de ne plus satisfaire les spécifications en durée de vie des équipements aéronautiques. Cette étude s'attache donc à montrer l'impact que peuvent avoir la réduction des dimensions ainsi que la nature des matériaux, sur la durée de vie des technologies numériques CMOS 500 à 45 nm. Pour cela, les mécanismes de défaillances du circuit intégré ont été modélisés et étudiés au travers de trois applications aéronautiques. En complément, des tests de vieillissement de type HCI et NBTI pratiqués sur une SRAM de génération 90 nm ont permis de valider les prédictions. Enfin, les travaux aboutissent à une méthodologie de sélection de composants fortement submicroniques pour une application spécifique, en fonction de la technologie.

Mots clés: électronique, circuit intégré, technologie, fiabilité, durée de vie, aéronautique, méthodologie

Reliability of CMOS Deep Sub Micron technologies (DSM) for civil avionic, space and military applications

Abstract

For some years, deep sub micron components have been used in aeronautic equipment for long term applications (generally few decades). However, these components are devoted to mass markets that are basically microcomputers and telecommunications and not especially to aeronautics. Hence, deep sub micron components are manufactured for high performance and low consumption needs and unfortunally, reliability is not the main concern. In order to supply these markets, a new generation of components generally arise every two years, introducing BEOL and FEOL scaling and/or new materials. As a consequence, these improvements have induced a lifetime degradation of devices that can threaten their use for longtime specifications of aeronautic equipement. This study points out the effects of scaling and material improvements on the lifetime degradation of CMOS integrated cicuits between the nodes 500-45 nm. To do so, silicon failure mechanisms have been modelised and investigated on three aeronautic applications. As a complement, some HCI and NBTI lifetests have been performed on 90 nm SRAM in order to validate previous lifetime predictions. Finally, this work leads to a methodology for the selection of deep sub micron components for a specific use, depending on the technology.

Key words: electronic, integrated circuit, technology, reliability, lifetime, aeronautic, methodology

