

1. 1. Introduction
2. 2. Registres protégés utilisables seulement en mode kernel
3. 3. Adressage de la mémoire
4. 4. Instructions protégées
5. 5. Cause d'entrée et sortie du noyau du système d'exploitation
 1. A. Entrée pour cause d'exceptions
 2. B. Entrée pour cause d'interruptions
 3. C. Entrée pour cause d'appels système
 4. D. Entrée pour cause de signal RESET
 5. E. Sortie du noyau
6. 6. Fonctionnement du registre d'état `c0_sr`
7. 7. Fonctionnement du registre de cause `c0_cause`

Documentation MIPS Architecture et assembleur (mode kernel)

1. Introduction

Ce document est la suite du document Documentation MIPS32 architecture et assembleur (mode user) (*Ce document est tiré du document initialement écrit par Alain Greiner*).

Le MIPS supporte deux modes de fonctionnement utilisateur (*user*) et système (*kernel*).

- Dans le mode *user*, certaines régions de la mémoire et certains registres du processeur sont protégés et donc inaccessibles. C'est dans ce mode que s'exécute les applications.
- Dans le mode *kernel*, toutes les ressources sont accessibles, c'est-à-dire toute la mémoire et tous les registres. Dans ce mode, toutes les instructions sont autorisées, à la fois les instructions standards (`add`, `or`, `lw`, `mul`, etc.), mais aussi les **instructions protégées** qui permettent de contrôler l'état de fonctionnement du processeur. C'est dans ce mode que s'exécute le noyau du système d'exploitation.

Ce document détaille les éléments de l'architecture externe du processeur et du langage d'assemblage spécifique au mode *kernel*.

2. Registres protégés utilisables seulement en mode kernel

En mode *kernel*, tous les registres sont accessibles, à la fois les registres non protégés et aussi les registres protégés. Pour rappel, les registres non protégés sont les GPR (`$0` à `$31`), le registre PC (accessible implicitement avec les instructions de branchement) et les registres HI et LO. Les registres protégés sont destinés au calcul alors que les **registres protégés** sont destinés au contrôle de l'état du processeur.

L'architecture du MIPS32 définit 32 registres protégés, numérotés de `$0` à `$31`, c'est-à-dire comme les registres GPR mais ils ont des instructions d'accès spécifiques. En effet, ces registres protégés ne sont accessibles que par des instructions protégées présentées dans la section 4.

Ces registres appartiennent au "coprocesseur système" n°0 (appelé aussi `C0` pour *Coprocessor 0*). Dans cette version du processeur MIPS32, il y en a 6. Ils sont tous utilisés par le système d'exploitation pour la gestion des interruptions, des exceptions et des appels système. Dans ce document, nous ferons précéder le numéro du registre protégé par `C0_` afin de ne pas les confondre avec les registres standards.

Le registre `C0_SR`

Le registre `SR` de `C0` est le registre d'état (*Status Register*). Il contient en particulier le bit qui définit le mode d'exécution du processeur: *user* ou *kernel*, ainsi que les bits de masquage des interruptions. Ce registre a le numéro `$12`.

Le registre `C0_CAUSE`

Le registre `CAUSE` de `C0` est le registre de cause (*Cause Register*). En cas d'interruption, d'exception ou d'appel système, le programme en cours d'exécution est dérivé vers le noyau du système d'exploitation. Le contenu de `C0_CAUSE` contient un code qui définit la cause d'appel du noyau. Ce registre a le numéro `$13`.

Le registre `C0_EPC`

Le registre `EPC` de `C0` est le registre d'exception (*Exception Program Counter*). Il contient : (i) soit l'adresse de retour (`PC + 4`) en cas d'interruption, (ii) soit l'adresse de l'instruction courante (`PC`) en cas d'exception ou d'appel système. Ce registre a le numéro `$14`.

Le registre `C0_BAR`

Le registre `BAR` de `C0` est le registre d'adresse illégale (*Bad Address Register*). En cas d'exception de type *adresse illégale*, il contient la valeur de l'adresse mal formée. Une adresse est illégale, par exemple, si vous tentez une lecture de mot (`lw`) à une adresse non-alignée (non multiple de 4) ou si vous tentez une lecture en dehors des segments d'adresse où se trouve de la mémoire. Ce registre a le numéro `$8`.

Le registre `C0_PROCID`

Le registre `PROCID` de `C0` est un registre en lecture seule contenant le numéro du processeur. Cet index « cablé » est utilisé par le noyau du système d'exploitation. Il n'a de sens que pour gérer des architectures multiprocesseurs (multicore). Ce registre possède le numéro `$15`.

Le registre `C0_COUNT`

Le registre `COUNT` de `C0` est le registre en lecture seule contenant le nombre de cycles exécutés depuis l'initialisation du processeur. Ce registre possède le numéro `$16`.

3. Adressage de la mémoire

L'espace d'adressage de la mémoire est découpé en 2 parties identifiées par le bit de poids fort de l'adresse (bit n°31). Quand le processeur est en mode *kernel* alors les 2 parties (protégée et non protégée) sont accessibles. Quand le processeur est en mode *user* alors seule la partie protégée est accessible.

Bit n°31 de l'adresse = 0	partie non protégée utilisable dans tous les modes du processeur destinée au programme de l'utilisateur
Bit n°31 de l'adresse = 1	partie protégée utilisable seulement en mode kernel réservée au noyau du système d'exploitation

Quand le processeur est en mode *user*, si une instruction essaie d'accéder à la mémoire avec une adresse de la partie *protégée* alors le processeur part en exception, c'est-à-dire que le programme fautif est dérouté vers le noyau du système d'exploitation.

4. Instructions protégées

La version du MIPS32 que nous utilisons possède une cinquantaine d'instructions, il y a les instructions standards utilisables quel que soit le mode d'exécution du processeur et il y a les instructions protégées qui ne sont utilisables qu'en mode *kernel*. Les instructions standards sont présentées dans le document sur [l'architecture et l'assembleur en mode user](#). Ce sont les instructions arithmétiques/logiques entre registres, les instructions de branchement, les instructions de lecture et écriture mémoire et l'instruction `syscall`. Nous utilisons 3 instructions protégées (utilisables seulement en mode *kernel*) : `mtc0`, `mfc0` et `eret`.

`mtc0` et `mfc0`

signifient respectivement *Move-To-Coprocessor-0* et *Move-From-Coprocessor-0*. Comme leur nom l'indique, elles permettent de déplacer le contenu des registres entre les bancs (GPR et Copro).

instruction assembleur	comportement dans le processeur	Remarques
<code>mtc0 \$GPR, \$C0</code>	COPRO.0 (\$C0) ? GPR (\$GPR)	\$C0 = \$8, \$12, \$13, \$14, \$15 OU \$16 \$GPR = \$0 ... \$31
<code>mfc0 \$GPR, \$C0</code>	GPR (\$GPR) ? COPRO.0 (\$C0)	\$C0 = \$8, \$12, \$13, \$14, \$15 OU \$16 \$GPR = \$0 ... \$31

`eret`

signifie *Exception-RETurn*, c'est-à-dire *retour d'une exception*. Nous allons voir en détail ce que cela signifie dans la section 5. Pour le moment, comprenez que c'est l'unique instruction permettant de sortir du mode *kernel* pour entrer ou retourner dans le mode *user*.

instruction assembleur	comportement dans le processeur	Remarques
<code>eret</code>	PC ? CO_EPC CO_SR.EXL ? 0	copie le contenu du registre CO_EPC (CO_\$14) dans le registre PC et met 0 dans le bit EXL du registre CO_SR (CO_\$12)

Codage des instructions protégées

Elles utilisent toutes le format R avec le champ `OPCOD` à la valeur `COPRO` (c.-à-d. `0b010000`). L'instruction est alors codée avec les bits 25 et 23 de l'instruction (ces deux bits sont dans le champs `RS`). Remarquez que `eret` à deux codages.

OPCOD	RS	RT	RD	SH	FUNC
-------	----	----	----	----	------

31 25 20 15 10 5 0

INS 23

INS 25 0 1

0 mfc0 mtc0 1 eret

Pour les instructions `mtc0` et `mfc0`, le premier argument est mis dans le champs RT et le second argument est mis dans le champs RD.

instruction	comportement	commentaire
<code>mtc0 RT, RD</code>	<code>C0_RD ? RT</code>	Recopie le contenu du registre GPR n°RT dans le registre du coprocesseur 0 n°RD
<code>mfc0 RT, RD</code>	<code>RT ? C0_RD</code>	Recopie le contenu du registre du coprocesseur 0 n°RD dans le registre GPR n°RT

Par exemple:

`mtc0 $5, $14` est codé avec `0x40857000`

En effet : `OPCODE = 010000`, le bit `INS 25` est à 0 et le bit `INS 23` est à 1, donc :

? `mtc0 $5, $14` = `0b010000|0.1..| $5 | $14 | .....| .....`

? `mtc0 $5, $14` = `0b010000|0.1..| 00101|01110| .....| .....`

? . est un joker qui peut être remplacé par 0 ou 1, utilisons 0

? `mtc0 $5, $14` = `0b010000|00100|00101|01110|00000|000000`

? `mtc0 $5, $14` = `0b0100 0000 1000 0101 0111 0000 0000 0000`

? `mtc0 $5, $14` = `0x40857000`

5. Cause d'entrée et sortie du noyau du système d'exploitation

Il existe quatre types d'évènements qui peuvent interrompre l'exécution "normale" d'un programme :

- les exceptions ;
- les interruptions ;
- les appels système (instructions `syscall`) ;
- et le signal RESET.

Dans tous ces cas, le principe général consiste à dérouter le programme vers un code spécial (appelée noyau du système d'exploitation) qui s'exécute en mode système et à qui il faut transmettre les informations minimales lui permettant de traiter le problème.

A. Entrée pour cause d'exceptions

Les exceptions sont des évènements "anormaux" détectés au moment de l'exécution des instructions. Ils sont le plus souvent liés à une erreur de programmation qui empêche l'exécution correcte de l'instruction en cours. La détection d'une exception entraîne l'arrêt immédiat de l'exécution de l'instruction fautive. Ainsi, on assure que l'instruction fautive ne modifie pas la valeur d'un registre visible ou de la mémoire. Les exceptions ne sont évidemment pas masquables, cela signifie que l'on ne peut pas interdire leur gestion. Il y a 7 types d'exception dans cette version du

processeur MIPS32 :

ADEL

Adresse illégale en lecture : adresse non alignée ou se trouvant dans la partie système alors que le processeur est en mode utilisateur.

ADES

Adresse illégale en écriture : adresse non alignée ou accès à une donnée dans la partie système alors que le processeur est en mode utilisateur.

DBE

Data Bus Erreur : le système mémoire signale une erreur en activant le signal BERR à la suite d'un accès de donnée à une adresse qui n'a pas de case mémoire associée. On dit qu'elle n'est pas *mappée*. Cette erreur est aussi nommée erreur de segmentation ('segmentation fault' en anglais).

IBE

Instruction Bus Erreur : le système mémoire signale une erreur en activant le signal BERR à l'occasion d'une lecture instruction. C'est le même problème que pour **DBE** mais cela concerne les instructions.

OVF

Dépassement de capacité : lors de l'exécution d'une instruction arithmétique (ADD ou ADDI), le résultat ne peut être représenté sur 32 bits. Par exemple, la somme de 2 nombres positifs donne un nombre négatif.

RI

OPCOD illégal : l'OPCOD ne correspond à aucune instruction connue, il s'agit probablement d'un branchement dans une zone mémoire ne contenant pas du code exécutable.

CPU

Coprocesseur inaccessible : tentative d'exécution d'une instruction privilégiée (`mtc0`, `mfc0`, `eret`) alors que le processeur est en mode utilisateur.

Dans tous les cas, le processeur doit passer en mode *kernel* et se brancher au noyau du système d'exploitation implanté à l'adresse `0x80000180`. Après avoir identifié que la cause est une exception (en examinant le contenu du registre `C0_CAUSE`), le noyau se branche alors au gestionnaire d'exception. Ici, toutes les exceptions sont fatales, il n'y a pas de reprise de l'exécution de l'application contenant l'instruction fautive. Le processeur doit cependant transmettre au gestionnaire d'exceptions l'adresse de l'instruction fautive et indiquer dans le registre de cause le type d'exception détectée. Lorsqu'il détecte une exception, le matériel doit donc:

- sauvegarder le registre PC (l'adresse de l'instruction fautive) dans le registre `C0_EPC` ;
- passer en mode *kernel* et masquer les interruptions en mettant 1 dans le bit EXL de `C0_SR` ;
- sauvegarder éventuellement l'adresse fautive dans `C0_BAR` ;
- écrire le type de l'exception dans le registre `C0_CAUSE` ;
- se brancher à l'adresse `0x80000180`.

B. Entrée pour cause d'interruptions

Dans un ordinateur, nous avons vu qu'il y a au moins un processeur, une mémoire et des contrôleurs de périphériques. Les périphériques permettent, par exemple, de communiquer avec le monde extérieur comme pour le terminal texte. Les périphériques reçoivent des commandes dans leur registres par des instructions de lecture et d'écriture (`lw/sw`) venant du processeur.

Lorsqu'ils ont terminées une commande ou lorsqu'ils ont reçus ou calculés des données, les contrôleurs de périphérique peuvent le signaler au processeur par des **requêtes d'interruption** (IRQ pour Interrupt Request en anglais). Une requête d'interruption est un signal d'état produit par un contrôleur de périphérique avec deux états possibles : **actif** (ou levé) qui signifie que le contrôleur demande que le noyau intervienne ou **inactif** (ou baissé) qui

signifie que le contrôleur n'a pas de demande. Les requête d'interruptions sont donc des notifications de fins de commandes, d'arrivée de données sur un canal d'entrée ou encore des tick d'horloge.

Les requêtes d'interruption sont envoyées par des *lignes d'interruption* qui relie les contrôleurs de périphériques au processeur. Une *ligne d'interruption* est un fil électrique qui peut prendre les deux états des requêtes : actif/inactif, ou levé/baissé.

Les requêtes d'interruption s'activent de manière asynchrone par rapport au programme en cours d'exécution sont des événements asynchrones provenant des contrôleurs de périphériques. Ces événements sont toujours attendus par le noyau du système d'exploitation.

Le processeur MIPS32 possède 6 entrées de lignes d'interruptions externes qui peuvent être masquées globalement ou individuellement. Masquer une interruption signifie ne pas en tenir compte. Nous n'utiliserons qu'une seule de ces 6 entrées dans le prototype des TP.

Les lignes d'interruption peuvent être actives ou inactives (on dit aussi levées ou baissées). L'activation d'une de ces lignes est une requête d'interruption provenant d'un contrôleur de périphérique.

Si elles ne sont pas masquées alors elles sont prises en compte à la fin de l'exécution de l'instruction en cours. Une requête émise par un contrôleur de périphérique doit être maintenue active par le contrôleur tant qu'elle n'a pas été prise en compte par le noyau du système d'exploitation.

Le processeur doit alors passer alors en mode système et se brancher au noyau du système d'exploitation. Après avoir identifié que la cause est une interruption (en examinant le contenu du registre `c0_cause`), le noyau se branche au gestionnaire d'interruption qui doit appeler une fonction appropriée pour le traitement de la requête. Cette fonction est appelée routine d'interruption ou ISR (pour *Interrupt Service Routine*). Comme il faut reprendre l'exécution de l'application en cours à la fin du traitement de l'interruption, il faut sauvegarder une adresse de retour. Lorsqu'il reçoit une requête d'interruption non masquée, le matériel doit donc :

- sauvegarder PC+4 (l'adresse de retour) dans le registre `c0_epc` ;
- passer en mode système et masquer les interruptions dans `c0_sr` ;
- écrire qu'il s'agit d'une interruption dans le registre `c0_cause` ;
- brancher à l'adresse `0x80000180program`.

En plus des 6 lignes d'interruption matérielles, le processeur MIPS32 possède un mécanisme d'interruption logicielle: Il existe 2 bits dans le registre de cause `c0_cause` qui peuvent être écrits par le logiciel au moyen de l'instruction privilégiée `mtc0`. La mise à 1 de ces bits déclenche le même traitement que les requêtes d'interruptions externes, s'ils ne sont pas masqués.

C. Entrée pour cause d'appels système

L'instruction `syscall` permet à une application de l'utilisateur de demander un service au noyau du système d'exploitation, comme par exemple effectuer une entrée-sortie. Le code définissant le type de service demandé au système, et d'éventuels paramètres doivent avoir été préalablement rangés dans des registres généraux. Quand le processeur exécute l'instruction `syscall`, il passe en mode '`kernel`' et se branche au noyau. Après avoir identifié que la cause est un appel système (en examinant le contenu du registre `c0_cause`), le noyau se branche au gestionnaire d'appels système. L'instruction `syscall` réalise les opérations suivantes :

- sauvegarder PC (l'adresse de l'instruction) dans le registre `c0_epc` (l'adresse de retour est `PC + 4`) ;
- passer en mode système et masquage des interruptions dans `c0_sr` : `c0_sr.EXL ? 1` ;
- écrire la cause du déroutement dans le registre `c0_cause` (ici `c0_cause.code ? 8`) ;
- brancher à l'adresse `0x80000180`.

D. Entrée pour cause de signal RESET

Le processeur possède également une entrée RESET dont l'activation pendant au moins un cycle entraîne le branchement inconditionnel du code de démarrage de l'ordinateur (code de boot). Ce code, implanté à l'adresse 0xBF C00000 doit principalement charger le code du noyau du système d'exploitation dans la mémoire depuis le disque ou le réseau et se brancher à la fonction d'initialisation du noyau. Cette fonction initialise les contrôleurs de périphériques et les structures internes du noyau et, à la fin elle se branche à la première application utilisateur. Dans notre modèle d'ordinateur, le noyau est pré-chargé en mémoire et le code de boot se contente d'appeler la fonction d'initialisation.

Cette requête est très semblable à une septième ligne d'interruption externe avec les différences importantes suivantes:

- elle n'est pas masquable :
- il n'est pas nécessaire de sauvegarder une adresse de retour.
- le gestionnaire de reset est implanté à l'adresse "0xBFC00000".

Dans ce cas, le processeur doit :

- passer en mode système et masquer les interruptions dans SR
- brancher à l'adresse "0xBFC00000"

E. Sortie du noyau

Avant de reprendre l'exécution de l'application qui a effectué un appel système (instructions `syscall`) ou qui a été interrompu par une interruption, il est nécessaire d'exécuter l'instruction `eret`. Cette instruction modifie le contenu du registre `c0_sr`, et effectue un branchement à l'adresse contenue dans le registre `C0_EPC`.

6. Fonctionnement du registre d'état c0_sr

Le registre `c0_sr` contient l'état du processeur. Cela concerne son comportement vis-à-vis des requêtes d'interruptions, c'est-à-dire les masques des interruptions matérielles et logicielles, et le mode d'exécution, mode système (*kernel*) ou en mode utilisateur (*user*).

- La figure suivante présente le contenu des 16 bits de poids faible du registre `c0_sr`. Cette version du MIPS32 n'utilise que 12 bits:

IM[7:0]	0	0	0	UM	0	ERL	EXL	IE
---------	---	---	---	----	---	-----	-----	----

15	8	7	6	5	4	3	2	1	0
----	---	---	---	---	---	---	---	---	---

		0 = toutes les interruptions sont masquées
IE	Interrupt Enable	1 = interruptions non-masquées mais elles peuvent l'être avec IM[7:0]

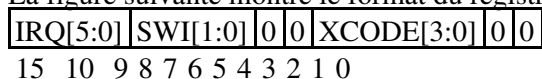
EXL	Exception Level	0 = aucun effet sur le processeur 1 = le processeur vient d'entrer dans le noyau et donc le processeur est en mode <i>kernel</i> et interruptions massquées
ERL	Error Level	1 = après le signal reset ou certaines erreurs de la mémoire
UM	User Mode	0 = mode d'exécution <i>kernel</i> 1 = mode d'exécution <i>user</i>
IM! [7:0]	Masques individuels	pour les six lignes d'interruption matérielles (bits IM[7:2]) et pour les 2 interruptions logicielles (bits IM[1:0])

- Quelques remarques :
 - ♦ Le processeur a le droit d'accéder aux ressources protégées (registres du coprocesseur 0 c0), et aux adresses mémoires $\geq 0x80000000$) si et seulement si le bit UM vaut 0, ou si l'un des deux bits ERL et EXL vaut 1.
 - ♦ Les interruptions sont autorisées si et seulement si le bit IE vaut 1, et si les deux bits ERL et EXL valent 00, et si le bit correspondant de IM vaut 1.
 - ♦ Les trois types d'événements qui déclenchent le branchement au noyau: (interruptions, exceptions et appels système) forcent le bit EXL à 1, ce qui masque les interruptions et autorise l'accès aux ressources protégées.
 - ♦ L'activation du signal RESET qui force le branchement au code de boot force le bit ERL à 1, ce qui masque les interruptions et autorise l'accès aux ressources protégées.
 - ♦ L'instruction `eret` force le bit EXL à 0.
- Lors de l'activation du RESET :
 - ♦ `c0_sr` contient donc la valeur 0x0004 (0b000000000000000100).
 - ♦ Pour exécuter un programme utilisateur en mode protégé, avec interruptions activées, il doit contenir la valeur 0xFF11.
 - ♦ Le noyau doit écrire la valeur 0xFF13 dans `c0_sr` et l'adresse de la première fonction du programme utilisateur dans `c0_epc` avant d'appeler l'instruction `eret`.

7. Fonctionnement du registre de cause `c0_cause`

Le registre `c0_cause` contient trois champs. Les 4 bits du champ `XCODE[3:0]` définissent la cause de l'appel du noyau. Les 6 bits du champ `IRQ[5:0]` représentent l'état des lignes d'interruption externes au moment de l'appel au noyau. Les 2 bits `SWI[1:0]` représentent les requêtes d'interruption logicielle.

- La figure suivante montre le format du registre de cause CR :



- Les valeurs possibles du champ `XCODE` sont les suivantes :

0000	INT	Interruption
0001		Inutilisé
0010		Inutilisé
0011		Inutilisé
0100	ADEL	Adresse illégale en lecture

0101	ADES	Adresse illégale en écriture
0110	IBE	Bus erreur sur accès instruction
0111	DBE	Bus erreur sur accès donnée
1000	SYS	Appel système (<code>syscall</code>)
1001	BP	Point d'arrêt (<code>break</code>)
1010	RI	OPCOD illégal
1011	CPU	Coprocasseur inaccessible
1100	OVF	Overflow arithmétique
1101		Inutilisé
1110		Inutilisé
1111		Inutilisé