# Exécution de plusieurs systèmes d'exploitation sur une puce manycore CC-Numa sécurisée

Jean-Baptiste BRÉJON

Encadrant : Quentin MEUNIER
Lip6 - SoC - ALSOC

10 Septembre 2015

# Contexte du stage

- Projet ANR TSUNAMY - LIP6 (Alsoc)
- Partenaires : Lab-STICC (Lorient), CEA LIST, Laboratoire Hubert Curien (St Etienne)
- Thèmes du projet
  - Architecture manycore
  - Cloud computing
  - Problématiques de sécurité
- Apports du LIP6
  - Architecture TSAR (CC-NUMA)
  - Système d'exploitation ALMOS

# Mise en perspective

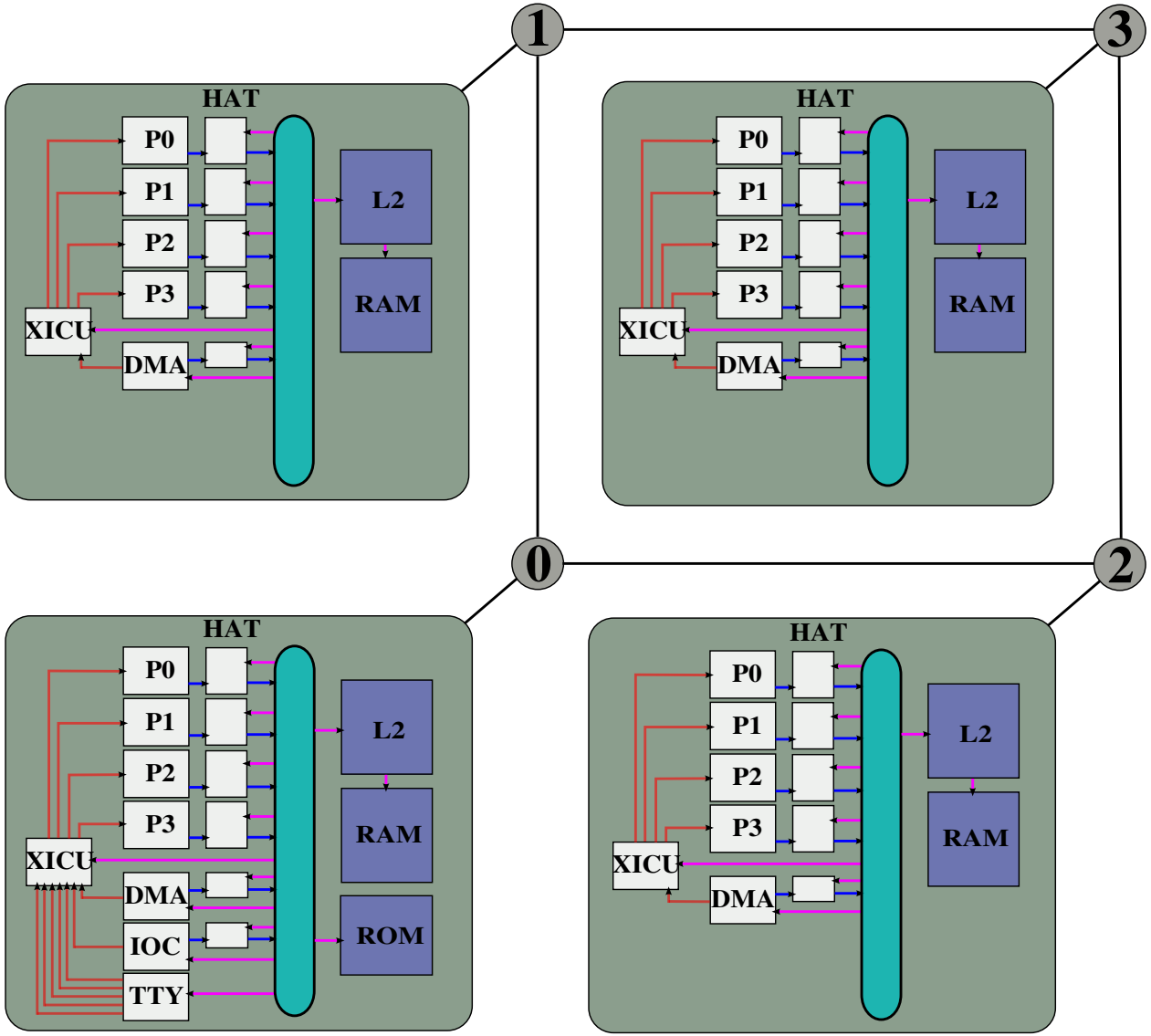
## Buts de Tsunamy

- Exécuter plusieurs systèmes d'exploitation
- Isolation des systèmes d'exploitation
- Chiffrement du code des systèmes d'exploitation
- Chiffrement et déchiffrement à la volée des données écrites ou lues sur le disque

## Apports du stage

- Exécuter plusieurs systèmes d'exploitation (hyperviseur "simple")

# Sujet : Démarrer 2 ALMOS sur une plateforme à 4 clusters



# Contexte

## Terminologie

- Système d'exploitation (OS) : ALMOS (utilisé dans le cadre de se stage).
- Instance d'un OS : Spécialisation d'un OS, qui comporte un système de fichiers, un code kernel binaire et un bootloader.
- Machine virtuelle (VM) : Sous-ensemble de clusters associé à une instance d'OS.

# Définition et analyse du problème

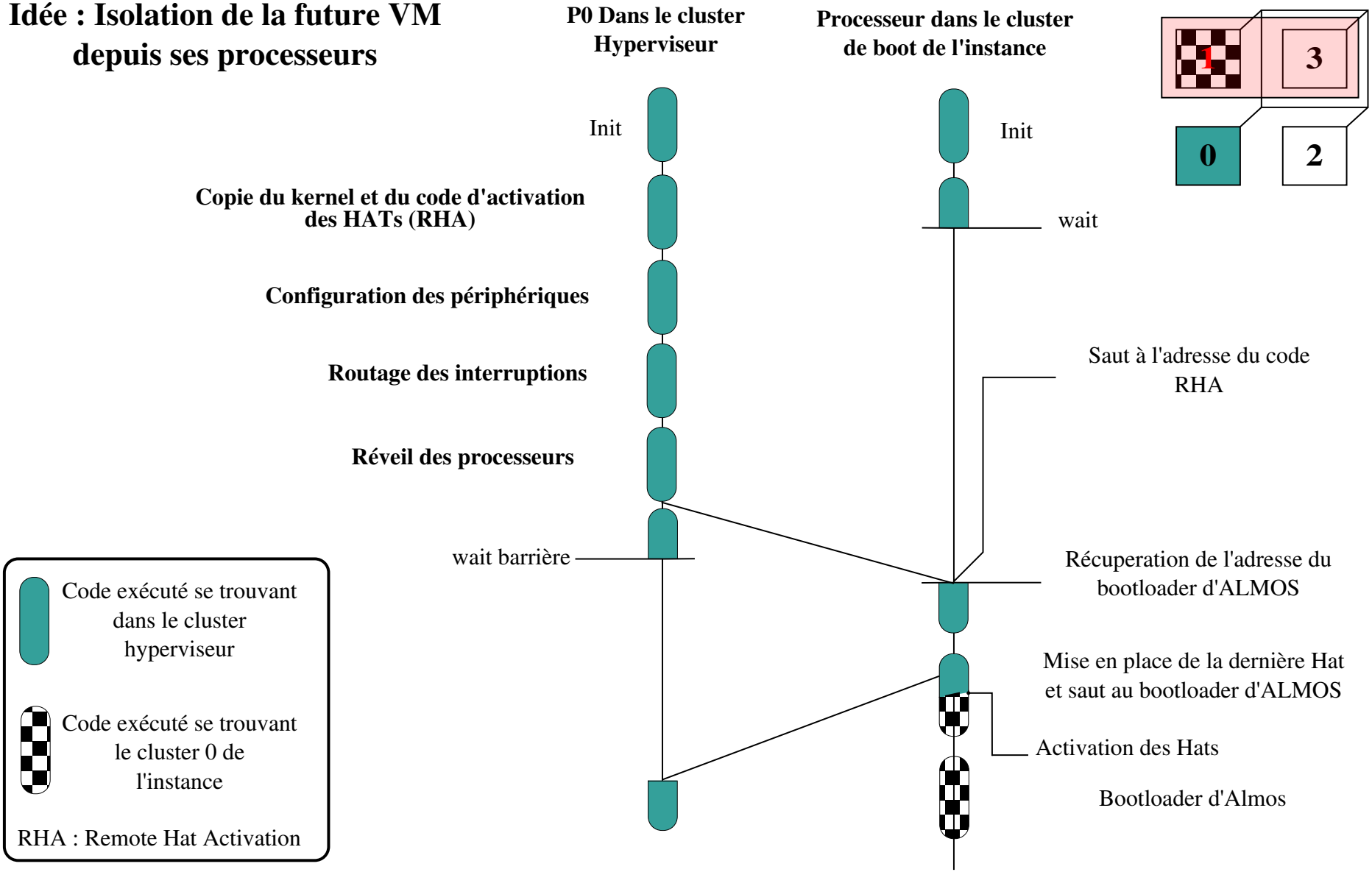
- Canaux des périphériques non répliqués
  - IOC : contrôleur de disque
    - Un seul canal
    - Un seul disque
  - MULTI_TTY : contrôleur de terminaux
- Routage des interruptions des périphériques non répliqués
  - Pas d'accès à l'intérieur de l'instance pour l'hyperviseur
- Démarrage des machines virtuelles
  - Isoler puis réveiller ?
  - Réveiller puis isoler ?
- Affichage des machines virtuelles sur les terminaux
  - Un canal = une fenêtre

# Solutions

- Routage des interruptions $\rightarrow$ Composant IOPIC
- Canaux de l'IOC $\rightarrow$ MULTI_IOC

# Solution pour le démarrage des machines virtuelles

**Idée : Isolation de la future VM depuis ses processeurs**



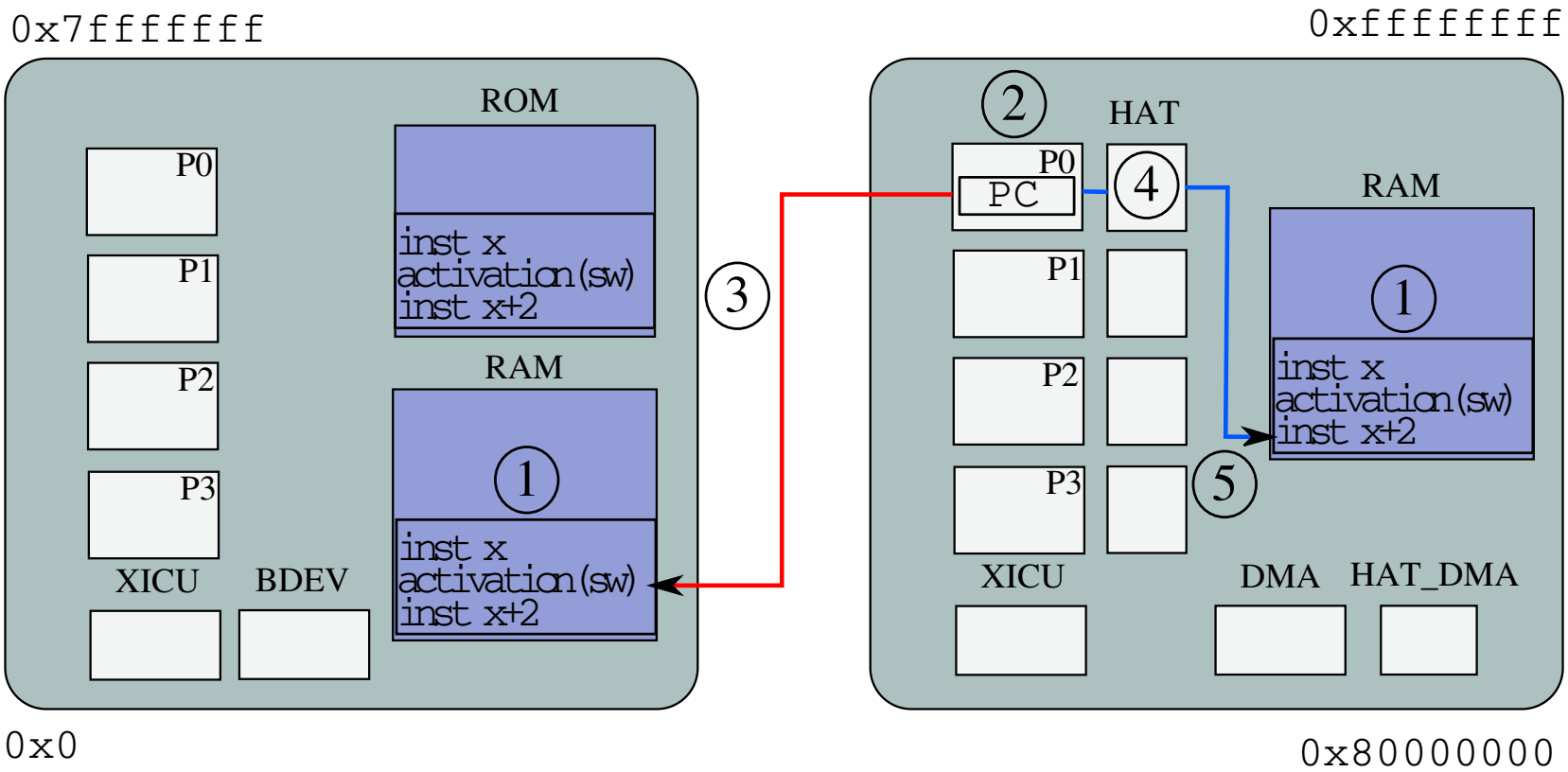
# Procédure de recette

1. Lancement d'un ALMOS sur 1 cluster sans IOPIC
2. Lancement d'un ALMOS sur 1 cluster avec IOPIC
3. Démarrage de deux ALMOS
   - Développement MULTI_IOC
4. Démarrage des VMs en utilisant le Shell hyperviseur
   - Développement MULTI_TTY_VT
5. Lancement de l'application "hello" sur les deux VMs

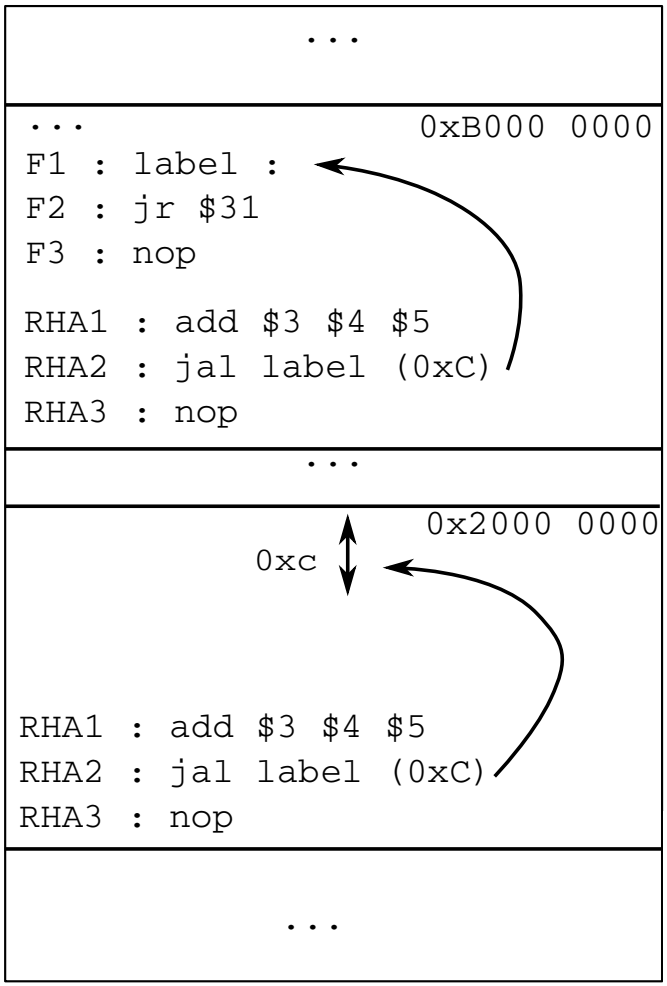
# Différence avec la spécification

- Ajout du composant MULTI_HAT devant le composant MULTI_IOC
- Procédure de démarrage d'une VM plus complexe

# Copie du code RHA dans la RAM

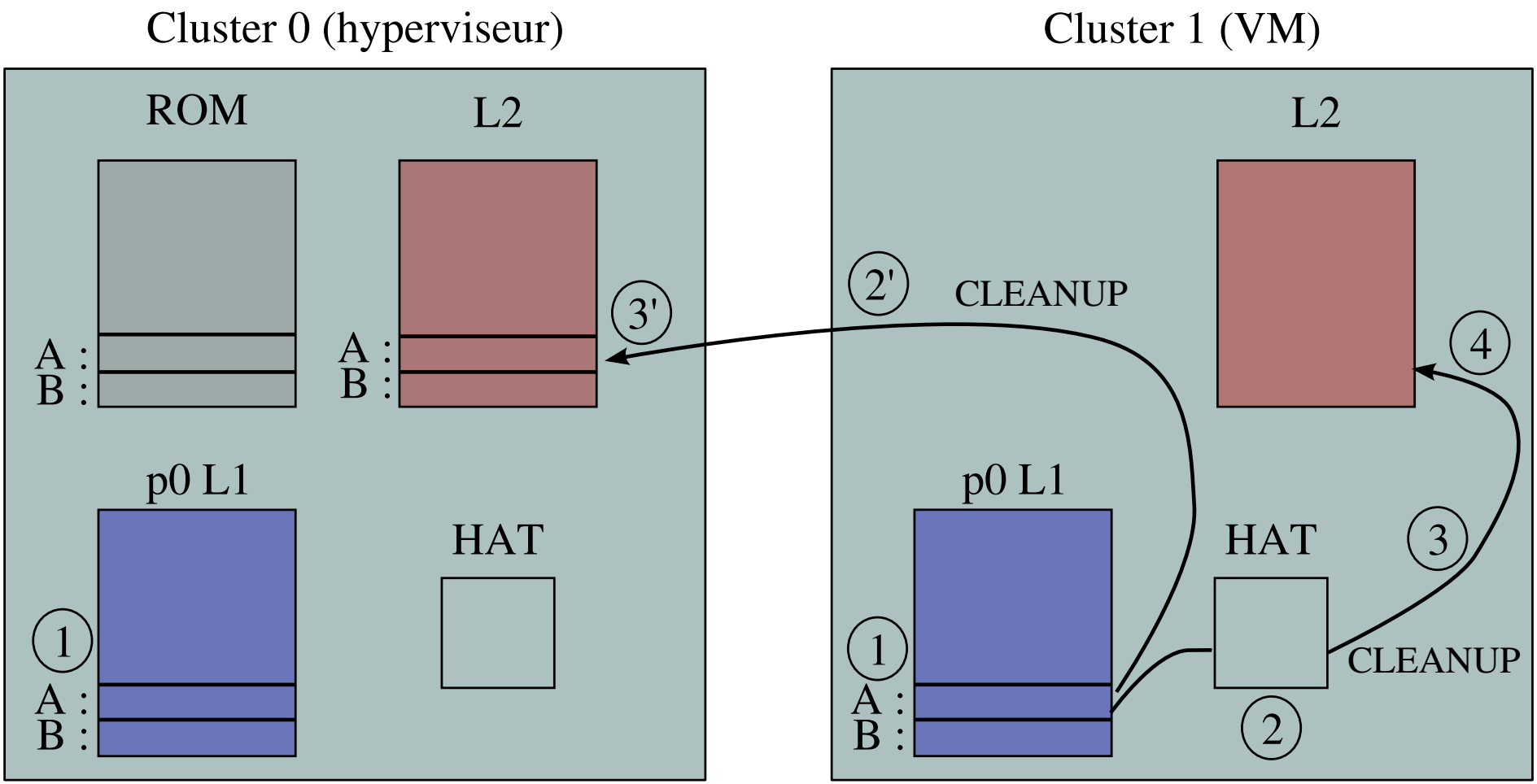


# Appel de fonctions dans le code RHA

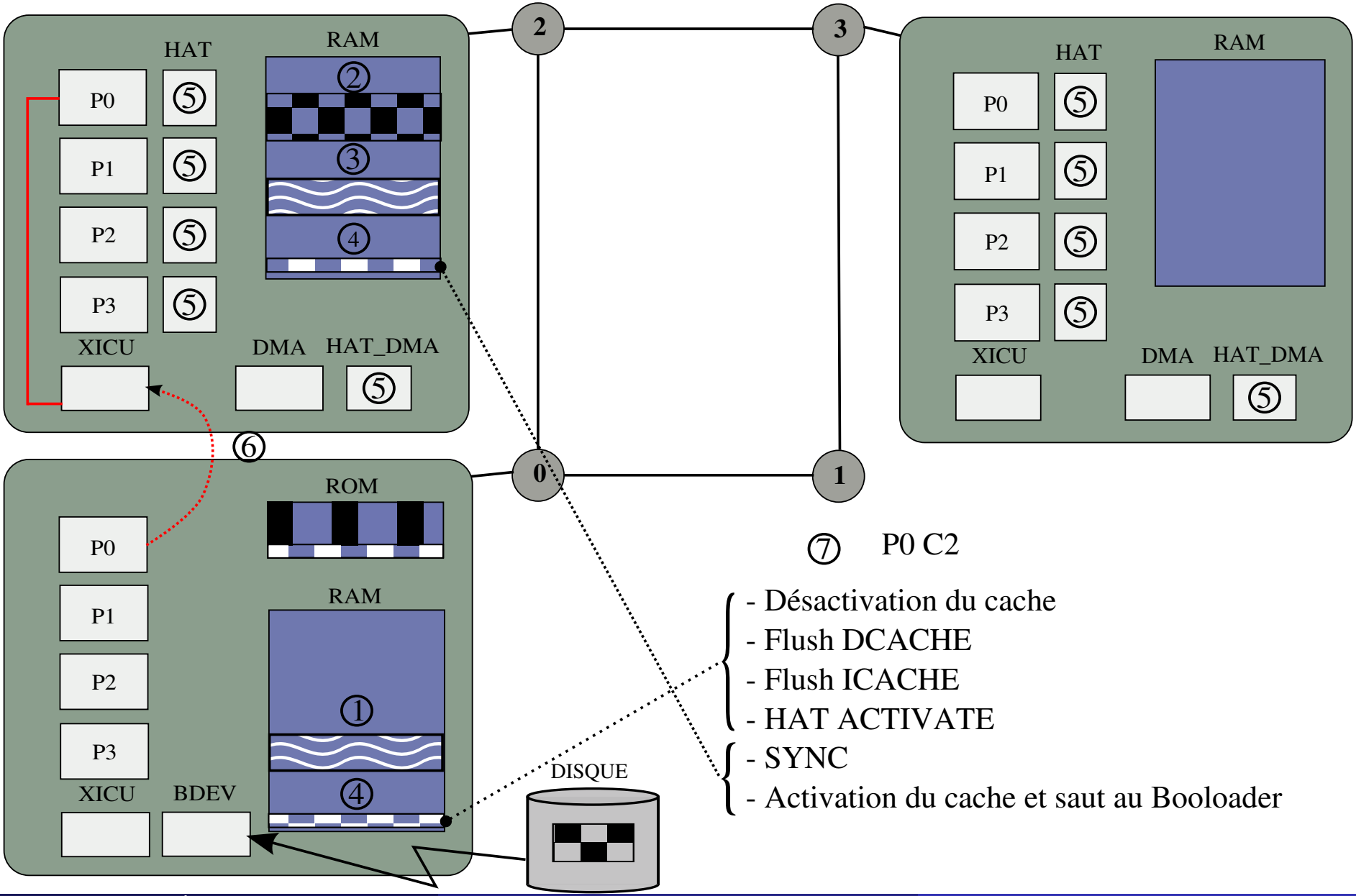



| | |
|---|---|
| JAL ADDR28 | $R_A = PC + 8; PC = PC_{31:28} :: ADDR28^{\varnothing}$ |
| JALR RD, Rs | $R_D = PC + 8; PC = Rs$ |

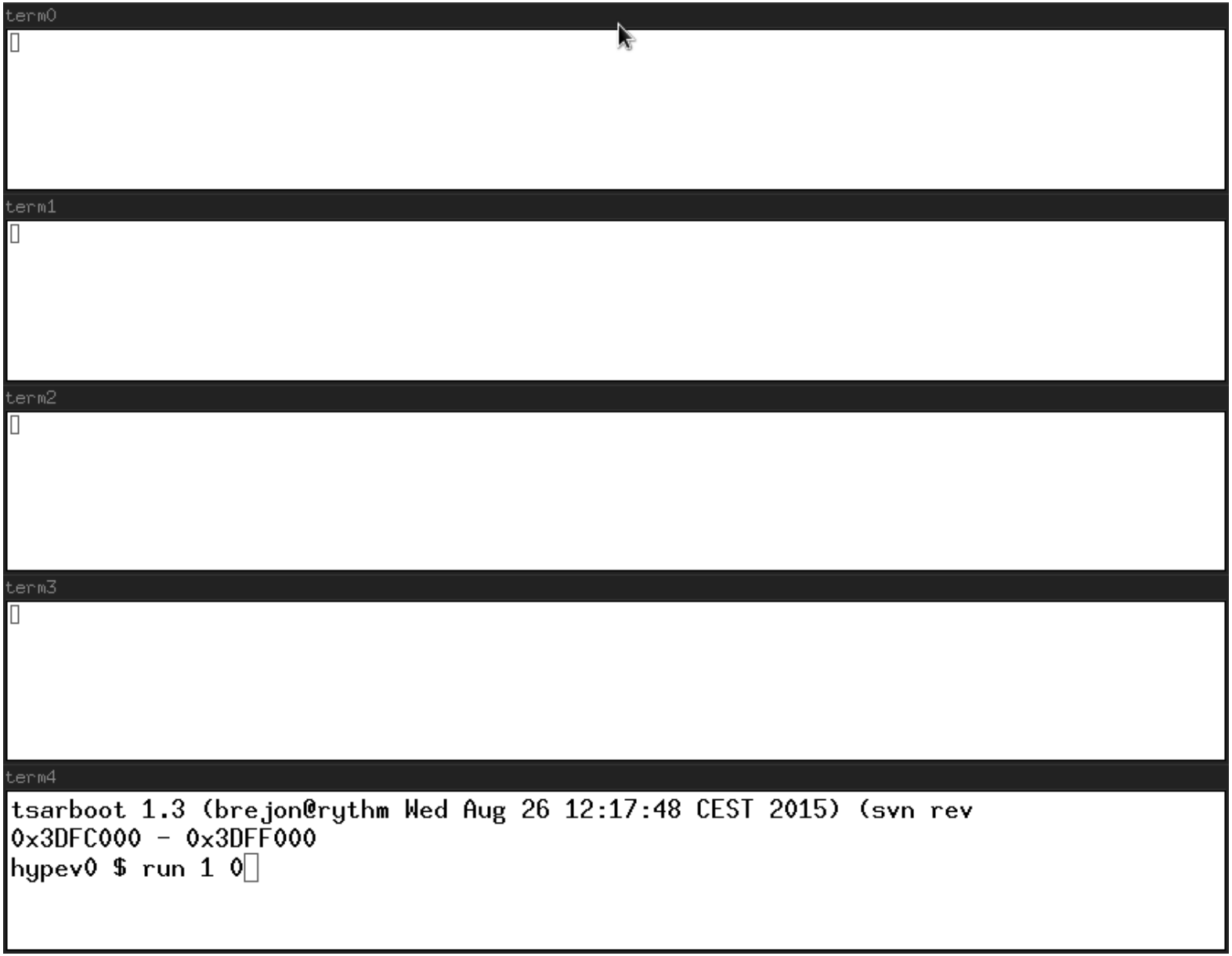
# Cohérence des caches



# Exemple complet : séquence de démarrage



# État initial



# Lancement de la première VM

term0_1.txt

```
Wake up all processors                    IPI SENDS                          [ addr 0x1f
f80004 jump at 0x100000 ]
IPI SENDS                          [ addr 0x1ff80008 jump at 0x100000 ]
IPI SENDS                          [ addr 0x1ff8000c jump at 0x100000 ]
QM: gid : 0 -- hw_id : 16 -- lid : 0 -- cid : 0 -- outirq : 0
QM: gid : 1 -- hw_id : 17 -- lid : 1 -- cid : 0 -- outirq : 4
QM: gid : 2 -- hw_id : 18 -- lid : 2 -- cid : 0 -- outirq : 8
QM: gid : 3 -- hw_id : 19 -- lid : 3 -- cid : 0 -- outirq : 12
cluster_init: cid 0: start_addr 0x200000, limit_addr 0x3e00000, start_vaddr 0x80000
000
PPM 0 [ 15360, 14938, 11950, 2240, 746 ]
PPM 0 pages_tbl [0, 1, 2, 0, 1, 2, 2, 1, 1, 28]
Mapping Region <0x1ff80000 - 0x1ff81000> -> <0xff800000 - 0xff801000>
Found Device: XICU (Interrupt Control Unite With Integrated RT-Timers)
Base <0xff800000> Size <0x1000> Irq <-1>
Mapping Region <0x1ff00000 - 0x1ff01000> -> <0xff801000 - 0xff802000>
```

term1_1.txt

```
7 found
[/]>exec hello
```

term2_1.txt

term3_1.txt

term4

```
launching almos 0 on 1
lp = 3, gp = 19, cxy_id 1, bootentry addr = 0x100000
hypev0 $ run 2 1
```

# Lancement de la seconde VM

term0_2.txt

```
        Kernel Entry Point @0x8000ca44

Initialization of 2 Online Clusters     [ 152259178 ]
Mapping Region <0x0 - 0x3e00000> : <0x80000000 - 0x83e00000>
Mapping Region <0x80000000 - 0x83e00000> -> <0x83e00000 - 0x87c00000>

Wake up all processors                  IPI SENDS                       [ addr 0x1ff80004 ju
mp at 0x100000 ]
IPI SENDS                       [ addr 0x1ff80008 jump at 0x100000 ]
IPI SENDS                       [ addr 0x1ff8000c jump at 0x100000 ]
IPI SENDS                       [ addr 0x9ff80000 jump at 0x100000 ]
IPI SENDS                       [ addr 0x9ff80004 jump at 0x100000 ]
IPI SENDS                       [ addr 0x9ff80008 jump at 0x100000 ]
IPI SENDS                       [ addr 0x9ff8000c jump at 0x100000 ]
QM: gid : 1 -- hw_id : 33 -- lid : 1 -- cid : 0 -- outirq : 4
QM: gid : 0 -- hw_id : 32 -- lid : 0 -- cid : 0 -- outirq : 0
QM: gid : 2 -- hw_id : 34 -- lid : 2 -- cid : 0 -- outirq : 8
QM: gid : 3 -- hw_id : 35 -- lid : 3 -- cid : 0 -- outirq : 12
QM: gid : 5 -- hw_id : 49 -- lid : 1 -- cid : 1 -- outirq : 4
QM: gid : 6 -- hw_id : 50 -- lid : 2 -- cid : 1 -- outirq : 8
QM: gid : 4 -- hw_id : 48 -- lid : 0 -- cid : 1 -- outirq : 0
QM: gid : 7 -- hw_id : 51 -- lid : 3 -- cid : 1 -- outirq : 12
cluster_init: cid 0: start_addr 0x200000, limit_addr 0x3e00000, start_vaddr 0x80000
```

term1_2.txt

```
7 found
[/]>exec hello
```

term2_2.txt

term3_2.txt

term4

```
launching almos 1 on 2
lp = 3, gp = 51, cxy_id 3, bootentry addr = 0x100000
hypev0 $
```

# Commande switch

term0_1.txt

```
QM: setting mask 0x4 for output irq 8
QM: setting mask 0x8 for output irq 12
All clusters have been Initialized [ 7354892 ]

 ____       ___       ___       ___   _______   ________
/    \     |   |     |   \     /   | /  ___  \ /  ____  |
/  __  \    | |       |   \___/   | |  /   \  | | /    |_/
/  /  \  \   | |       | _       _ | | |     | | | |______
/  /____\  \  | |       | | \   / | | | |     | | \______  \
/   ______   \  | |     _ | |  \_/  | | | |     | |  _     | |
/   /      \   \  | |____/ | | |       | | |  \___/  | | \____/ |
/_____/    \_____\|_________/|___|     |___|  \_______/  |________/

                Advanced Locality Management Operating System
                  UPMC/LIP6/SoC (06 August 2015 - 14:22:58)

INFO: Building Distributed Quaternary Decision Tree (DQDT)
INFO: DQDT has been built [581189]
INFO: kernel replication started on cid 0 [7982045]
INFO: kernel replication done on cid 0 [3176]
INFO: Starting Thread Idle On Core 0    OK
INFO: Starting Thread Idle On Core 2    OK
```

term1_1.txt

```
[/BIN]>Placement Decision: Static 1-to-1
pid 0, tid 1, arg 0: Hello World !!
```

term2_1.txt

```
LibC: main ended
LibC: calling gomp tream_destructor
```

term3_1.txt

term4

```
hypev0 $ switch_all 1
vm terms showing : 1
hypev0 $
```

# Commande switch

term0_1.txt

```
QM: setting mask 0x4 for output irq 8
QM: setting mask 0x8 for output irq 12
All clusters have been Initialized [ 7354892 ]

 ____       ___       ___       ___   _______   ________
/    \     |   |     |   \     /   | /  ___  \ /  ____  |
/  __  \    | |       |   \___/   | |  /   \  | | /    |_/
/  /  \  \   | |       | _       _ | | |     | | | |______
/  /____\  \  | |       | | \   / | | | |     | | \______  \
/   ______   \  | |     _ | |  \_/  | | | |     | |  _     | |
/   /      \   \  | |____/ | | |       | | |  \___/  | | \____/ |
/_____/    \_____\|_________/|___|     |___|  \_______/  |________/

                Advanced Locality Management Operating System
                  UPMC/LIP6/SoC (06 August 2015 - 14:22:58)

INFO: Building Distributed Quaternary Decision Tree (DQDT)
INFO: DQDT has been built [581189]
INFO: kernel replication started on cid 0 [7982045]
INFO: kernel replication done on cid 0 [3176]
INFO: Starting Thread Idle On Core 0    OK
INFO: Starting Thread Idle On Core 2    OK
```

term1_1.txt

```
[/BIN]>Placement Decision: Static 1-to-1
pid 0, tid 1, arg 0: Hello World !!
```

term2_1.txt

```
LibC: main ended
LibC: calling gomp tream_destructor
```

term3_1.txt

term4

```
hypev0 $ switch_all 1
vm terms showing : 1
hypev0 $
```

# Conclusion et perspectives

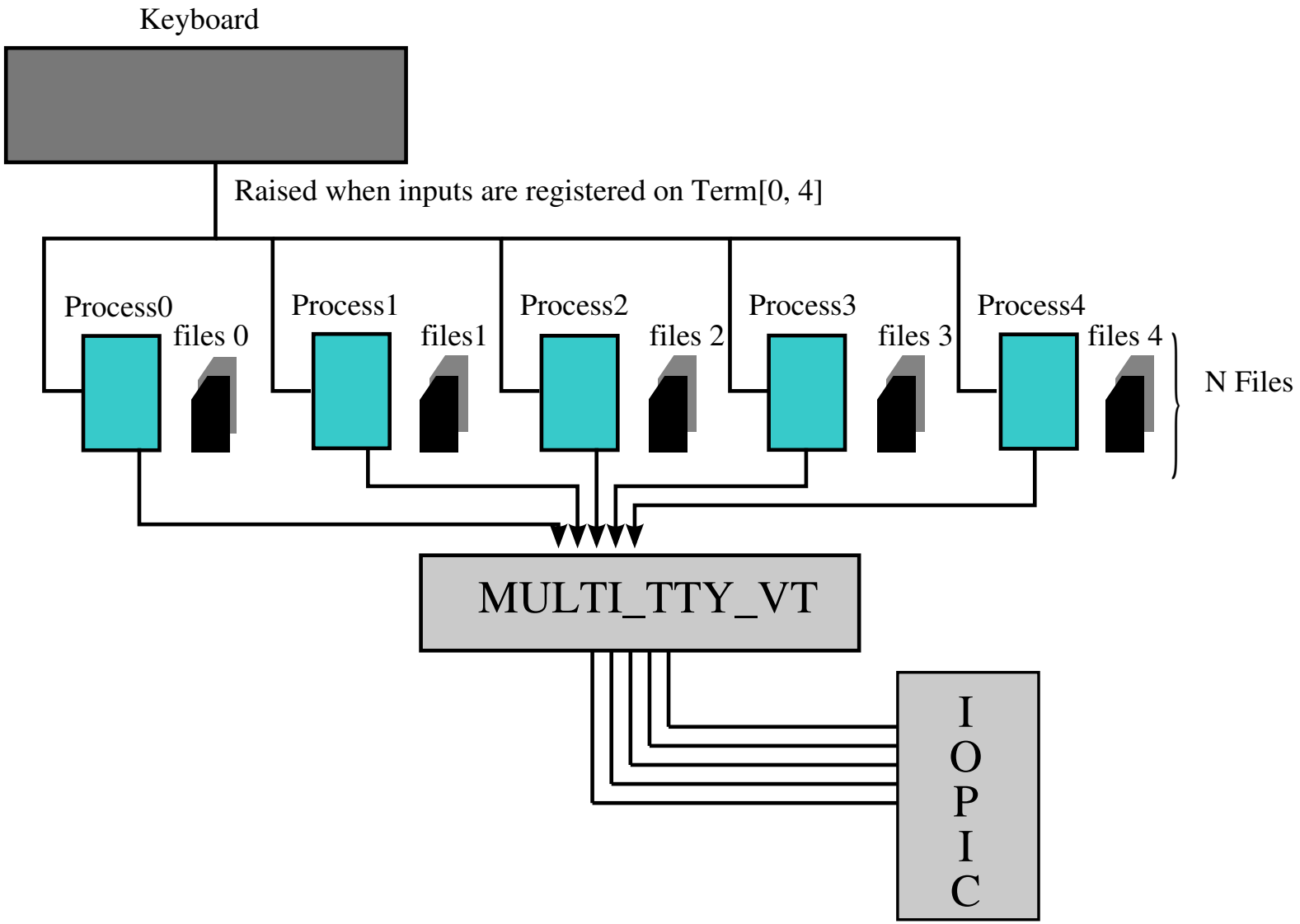
## Conclusion

- L'hyperviseur, couplé aux composants, permet de lancer plusieurs VMs sur la même machine
- Les VMs sont isolées entres elles et ne peuvent pas accéder à l'hyperviseur
- Migration vers une plateforme plus réaliste (IOB 40bits)

## Perspectives

- Destruction des VMs
- Isolation complète des VMs

# Solution pour affichage des machines virtuelles - Lecture



# Solution pour affichage des machines virtuelles - Écriture