

# 準仮想化ドメインのための PCI Hot-Plug

岩松洋介

NEC Corporation

# 目次

- 背景と動機
- 準仮想化ドメインのPCI Hot-Plugの実装
- 現状と今後の展望

# 背景と動機

# PCIパススルー

- PCIパススルーとは？
  - ゲストVMが物理PCIデバイスに直接アクセスする仕組み
- 利点
  - IO性能向上
    - IOの仮想化のオーバーヘッドを極小化
  - VMスケールABILITY向上
    - dom0のIOボトルネックを回避

Xen Summit Tokyo 2008

# PCIパススルーの欠点

- VMはデバイスに束縛される
  - 物理デバイスの状態は保存・復元できない
  - そのため、デバイスが割り当てられたVMのライブマイグレーションはできない
  - PCIパススルーは仮想化環境の柔軟性を損なう
- デバイスはVMに束縛される
  - 一つのVMがあるデバイスを排他的に専有
  - VM稼働中は他のVMからデバイスを使用できない
  - PCIパススルーはIOデバイスの不足につながる

Xen Summit Tokyo 2008

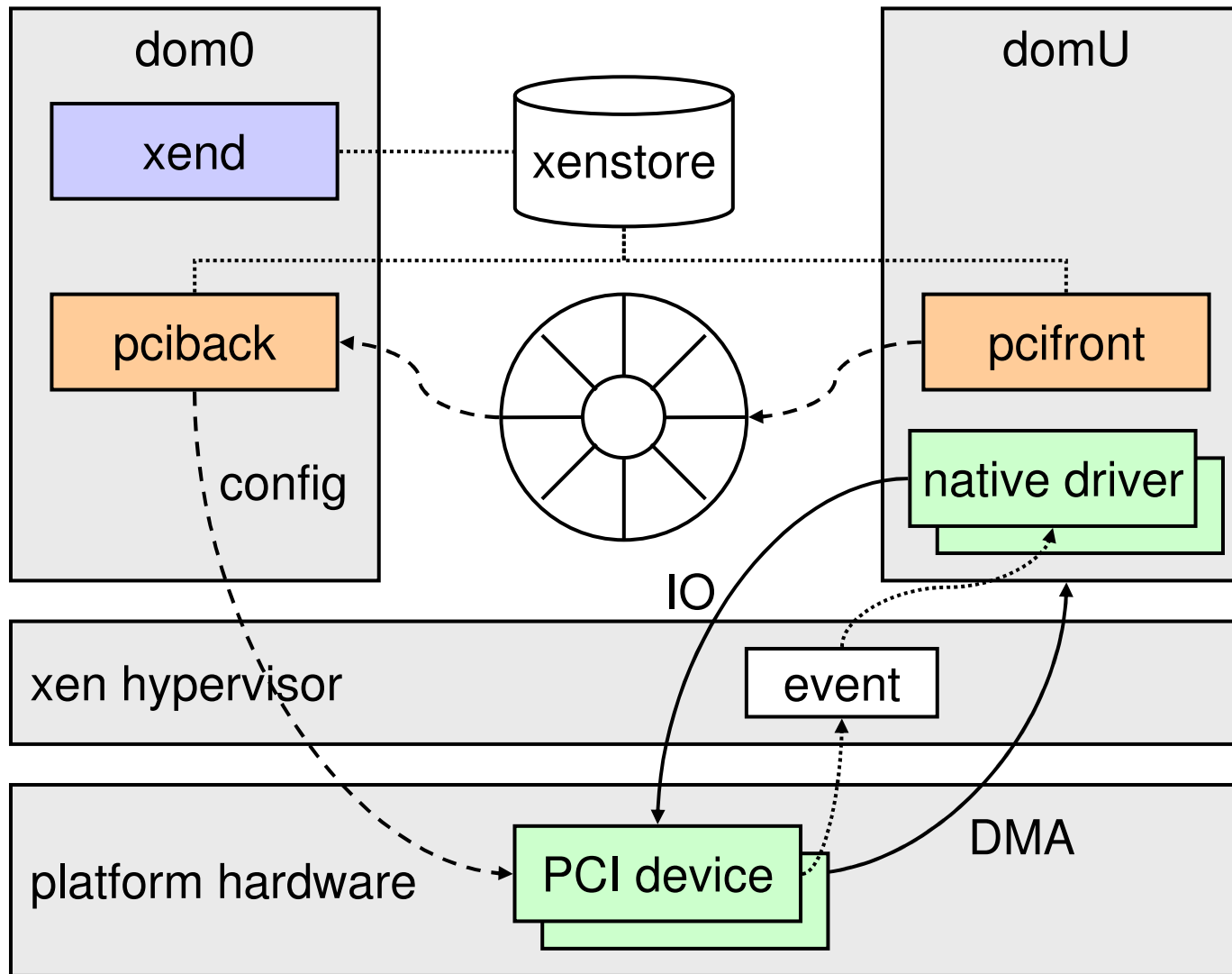
# VMに対するPCI Hot-Plug

- VMに対するPCI Hot-Plugとは？
  - 稼働中のVMに対し、物理PCIデバイスを追加・削除する仕組み
- 利点
  - VMの移動性向上
    - 割り当てられたデバイスを全て動的に削除した後、別マシンへのライブマイグレーションを実行
  - デバイスの利用効率向上
    - IO負荷の高いVMに対し、必要に応じてデバイスを動的に追加

Xen Summit Tokyo 2008

# 準仮想化ドメインの PCI Hot-Plugの実装

# 準仮想化のPCIパススルー (ドライバドメイン) の概要



Xen Summit Tokyo 2008

# 実装方針

- Xenbusの状態遷移

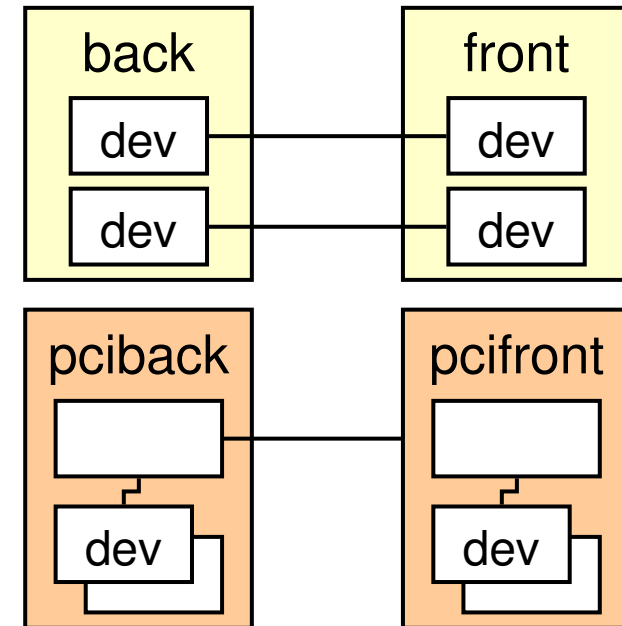
- 通常のPVドライバの場合

- 各々の仮想デバイスは独立したxenstoreノードとリング接続を持つ
    - 動的追加時に接続を開き、動的削除時に接続を閉じる

- pcifront・pcibackの場合

- ゲスト仮想PCIプラットフォームに対し接続は一つ
    - 各々のデバイスは属性として管理される

-> 各々のデバイスにサブ状態を導入し、接続を保ったままドライバを再設定



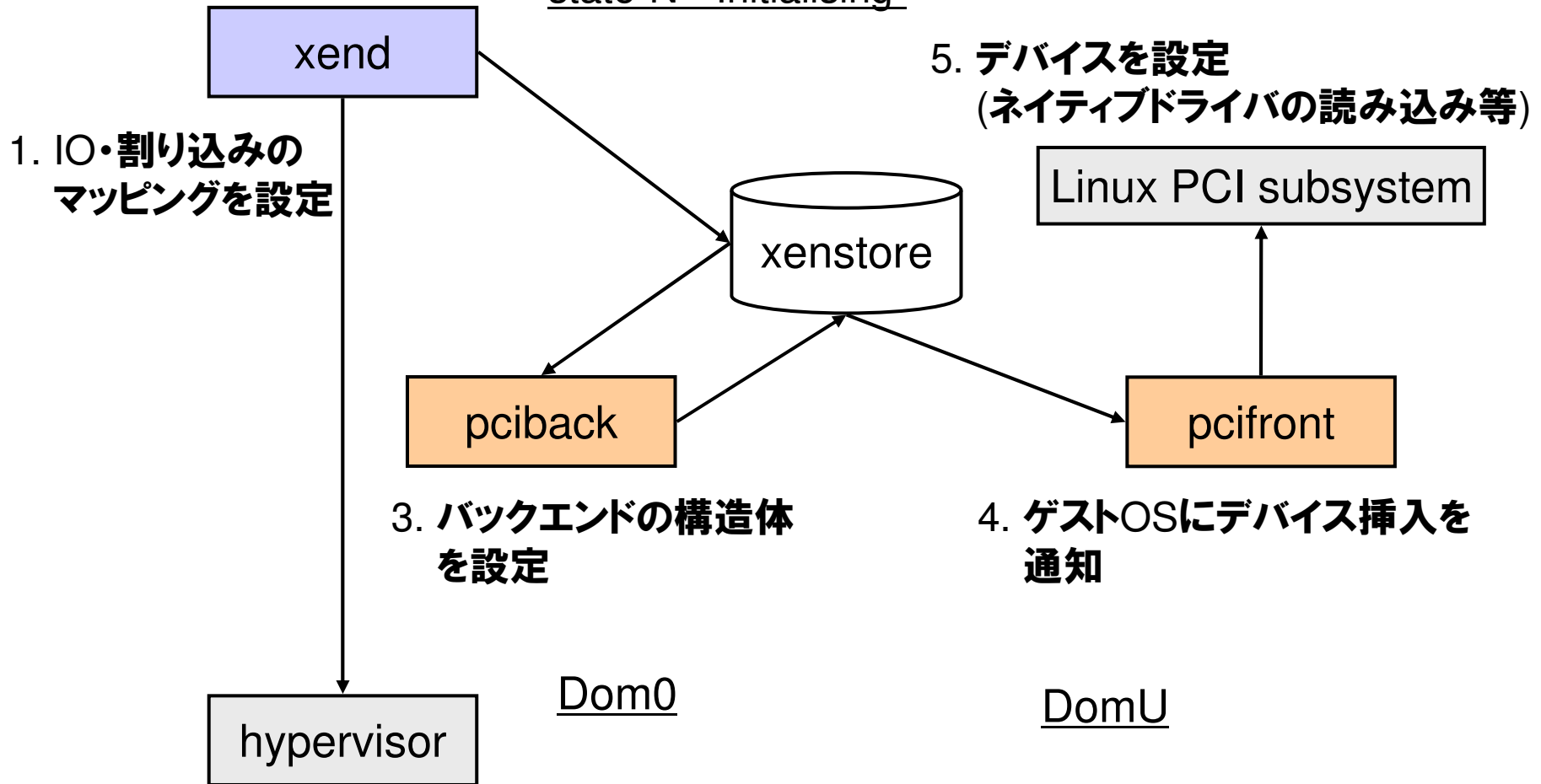
```
backend/pci/domid/0/  
state -> global state  
...  
dev-N -> phys B:D:F  
vdev-N -> virt B:D:F  
state-N -> sub-state  
...
```

Xen Summit Tokyo 2008

# Hot-Add

2. 追加するデバイスを指示し  
再設定を開始

state-N="Initialising"

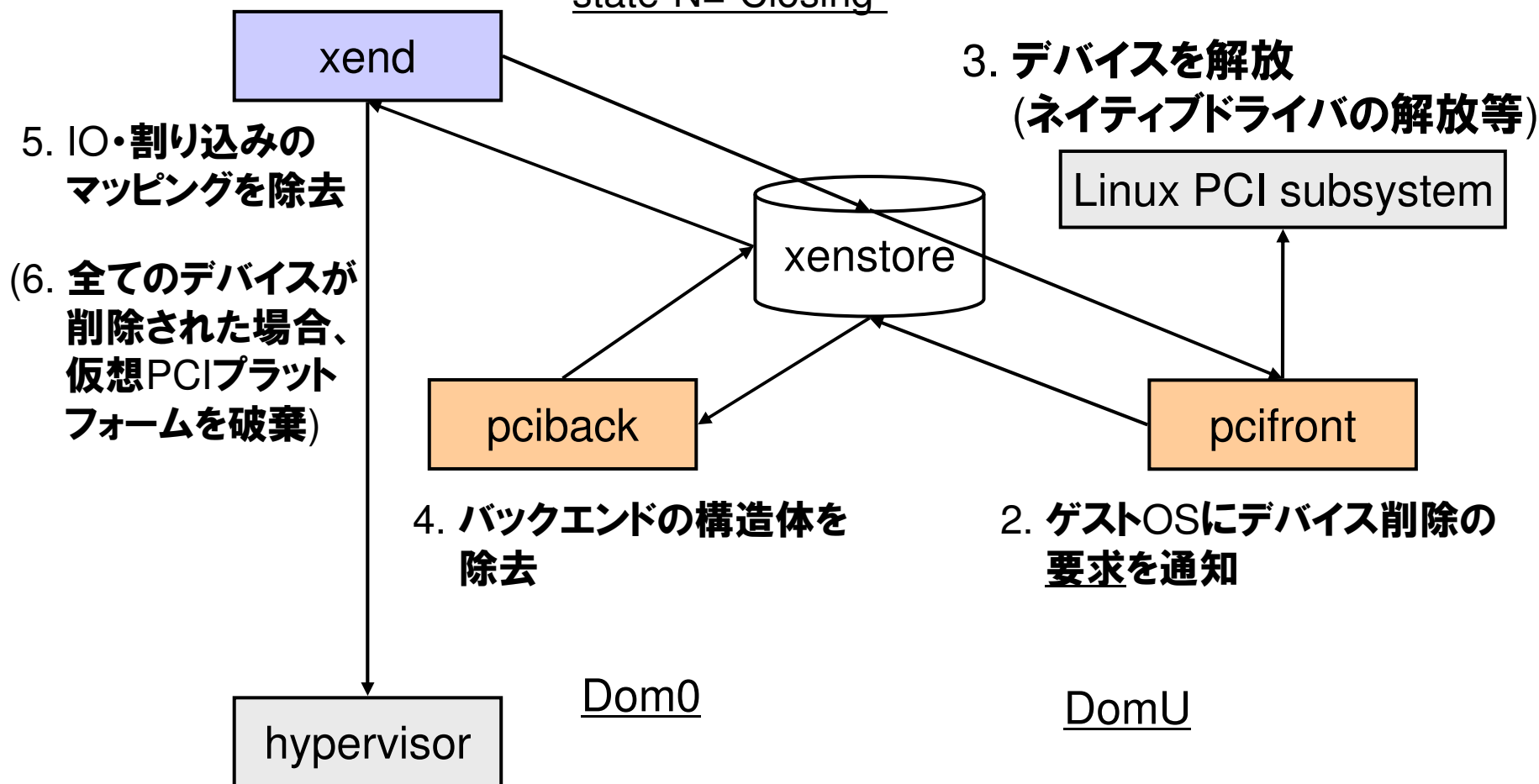


Xen Summit Tokyo 2008

# Hot-Remove

1. 削除するデバイスを指示し、再設定を開始

state-N="Closing"



Xen Summit Tokyo 2008

# 現状と今後の展望

# 現状と今後の課題

- Xen-3.3へ採用済みの項目
  - 準仮想化ドメインのPCI Hot-Plug
  - 完全仮想化ドメインのPCI Hot-Plug (Intelが開発)
    - コマンドは統一(xm pci-attach, pci-detach, pci-list)
  - Xen管理API経由のPCIパススルー制御も可能
- 今後の課題
  - 対応デバイスを増やす
  - スタブドメインのPCI Hot-Plug

## まとめ

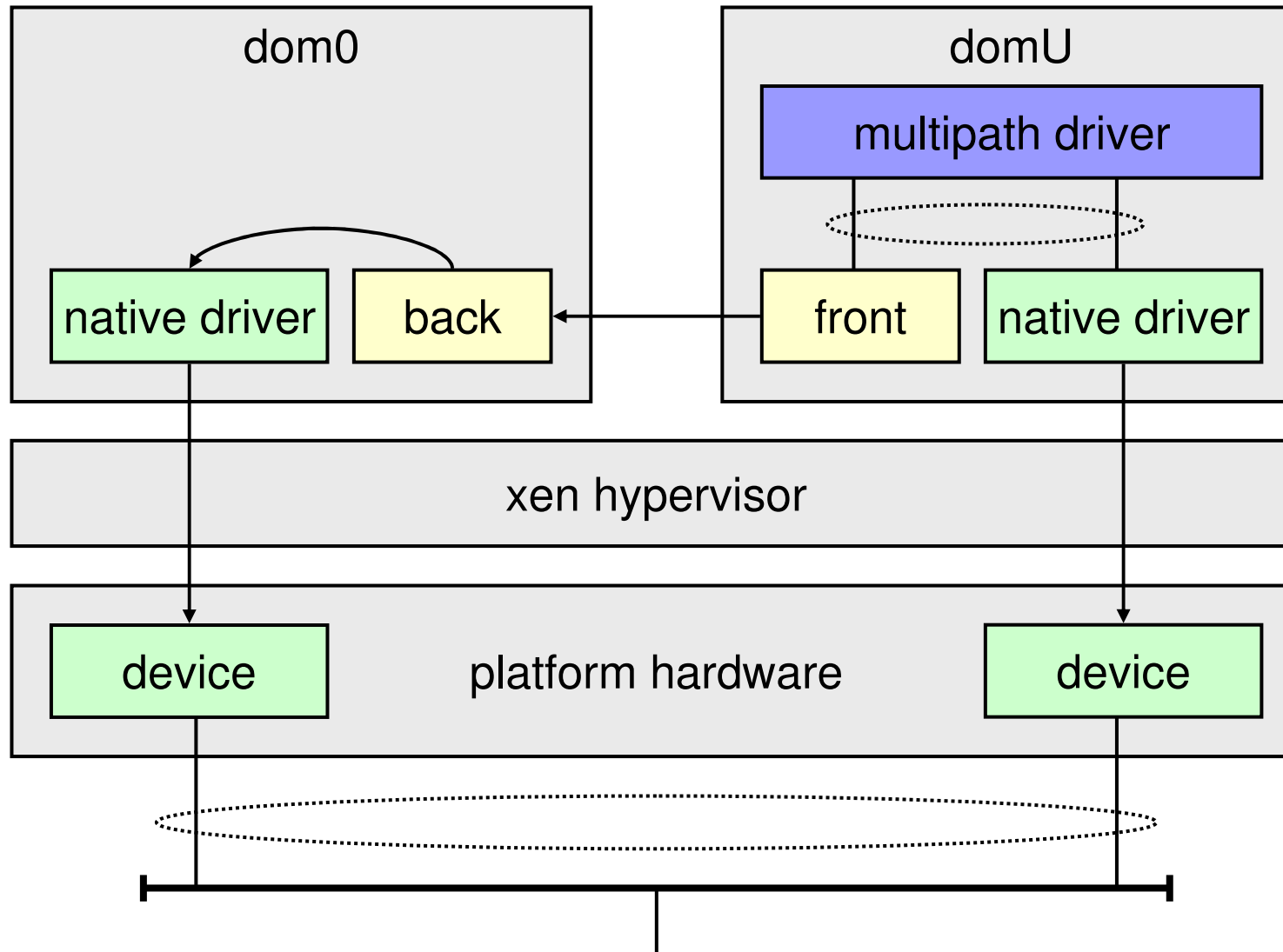
- VMへのPCI Hot-Plugにより、PCIパススルーを使用する仮想化システムにおいて、VMの移動性向上、デバイスの利用効率向上が期待できる
- 準仮想化ドメインのためのPCI Hot-Plugは、基本部分が実装されXen-3.3に取り込まれている

Empowered by Innovation

**NEC**

# 予備スライド

# 物理デバイスと仮想デバイスの結合

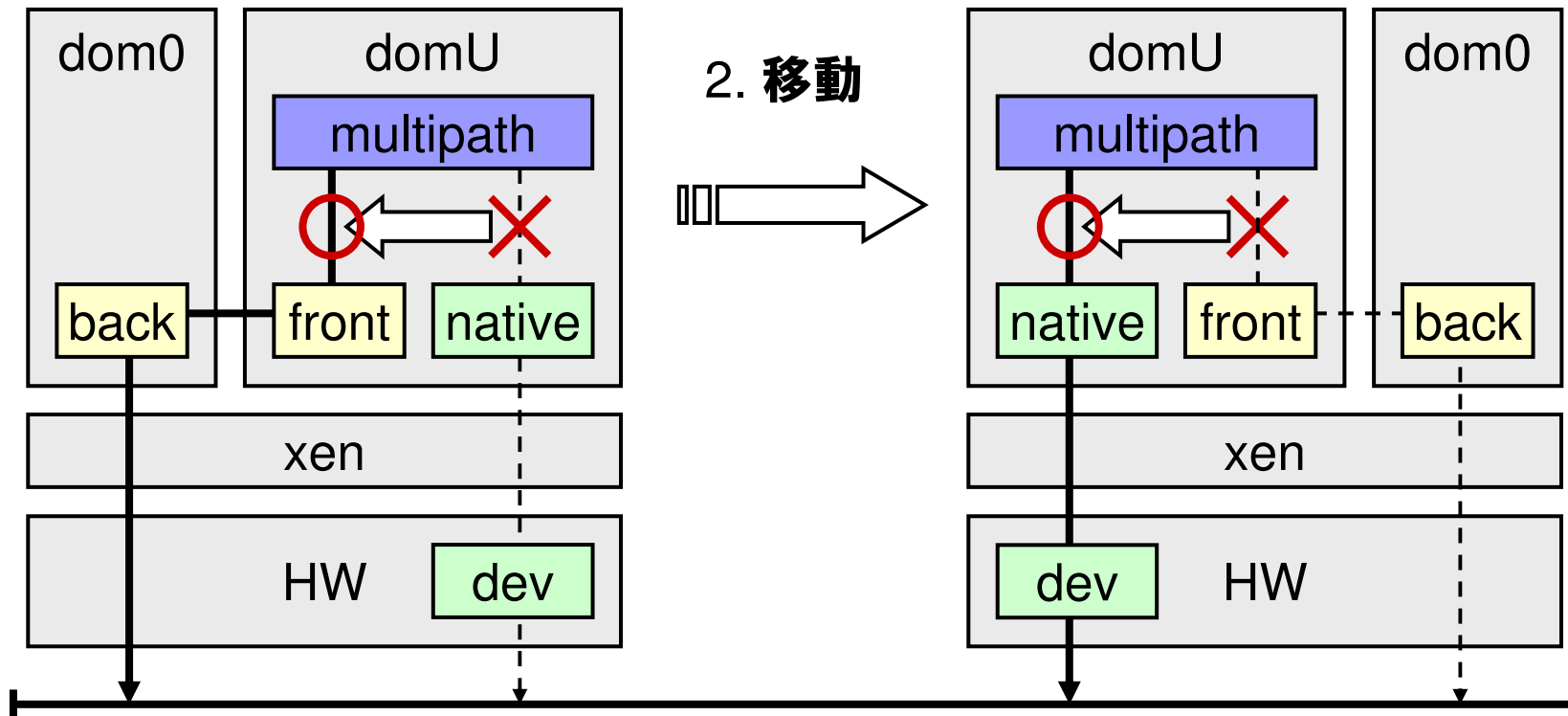


Xen Summit Tokyo 2008

# デバイスの割り当てられたVMのマイグレーション

1. 物理デバイスを動的に削除し、仮想パスに切り替え

3. 物理デバイスを動的に追加し、物理パスに切り替え



Xen Summit Tokyo 2008