

# PCI Hot-Plug for PV Domains

**Yosuke Iwamatsu**  
**NEC Corporation**

# Outline

- Background and Motivation
- Implementing PCI Hot-plug for PV Domains
- Current Status and Future Prospects

# Background and Motivation

# PCI Pass-through

- What is PCI pass-through ?
  - A guest VM directly accesses physical PCI devices
- Advantages
  - Improve IO performance
    - Minimize IO virtualization overhead
  - Improve VM scalability
    - Avoid IO bottleneck in dom0

Xen Summit Tokyo 2008

# Disadvantages of PCI Pass-through

- A VM is bound to a device
  - We cannot {save, restore} the state of a physical device
  - Thus, we cannot live-migrate a VM with assigned devices
  - PCI pass-through may sacrifice flexibility of the virtualized environment
- A device is bound to a VM
  - A device is exclusively used by a single VM
  - Other VMs cannot use the device as long as the owner VM is running
  - PCI pass-through may result in a shortage of IO devices

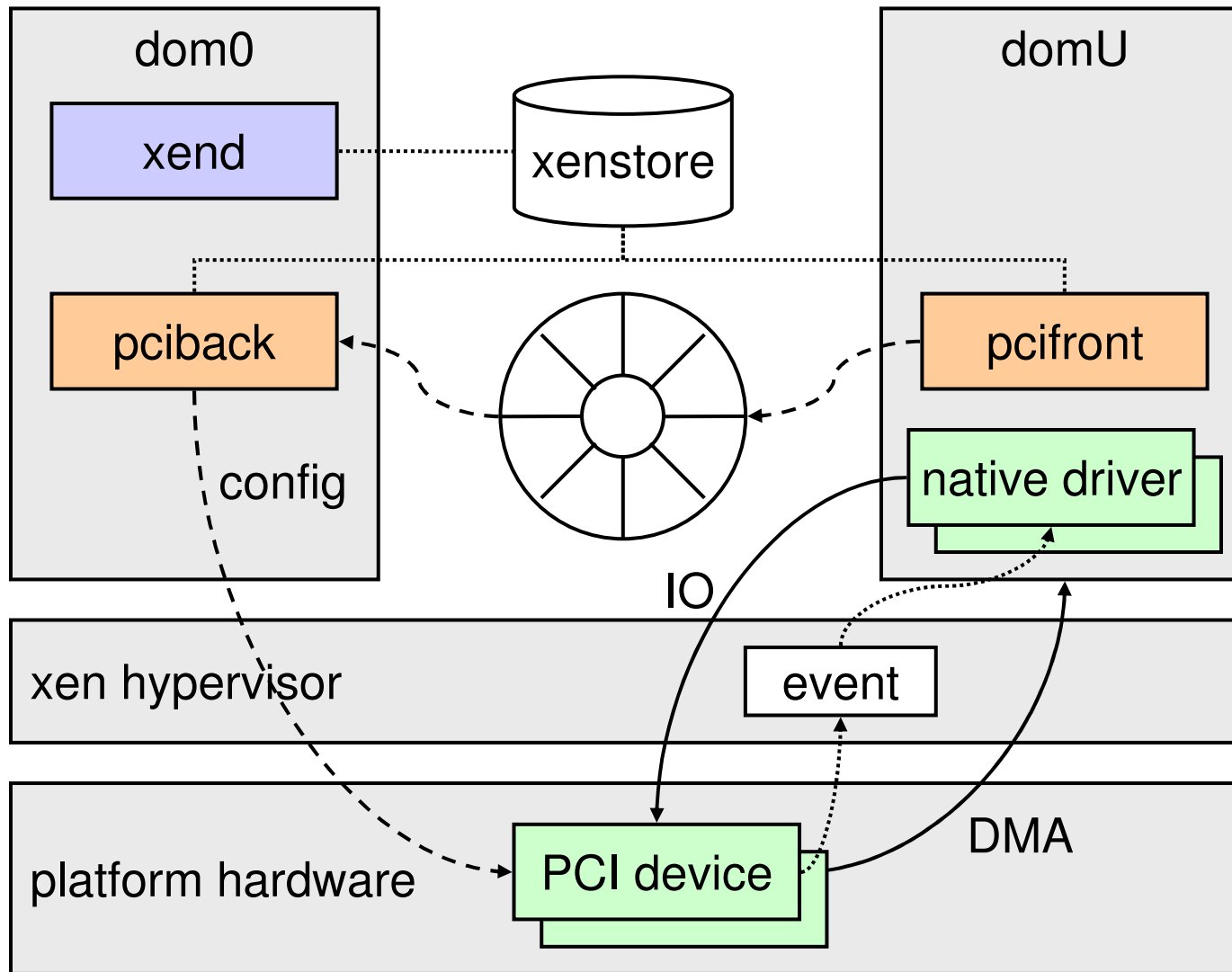
# PCI Hot-Plug for VM

- What is PCI hot-plug for VM ?
  - To {add, remove} a physical PCI device {to, from} a running VM
- Advantages
  - Improve VM mobility
    - Hot-remove all the assigned devices and live-migrate to another machine
  - Improve device utilization efficiency
    - Hot-add a device to a VM under IO-intensive workloads on demand

Xen Summit Tokyo 2008

# Implementing PCI Hot-Plug for PV Domains

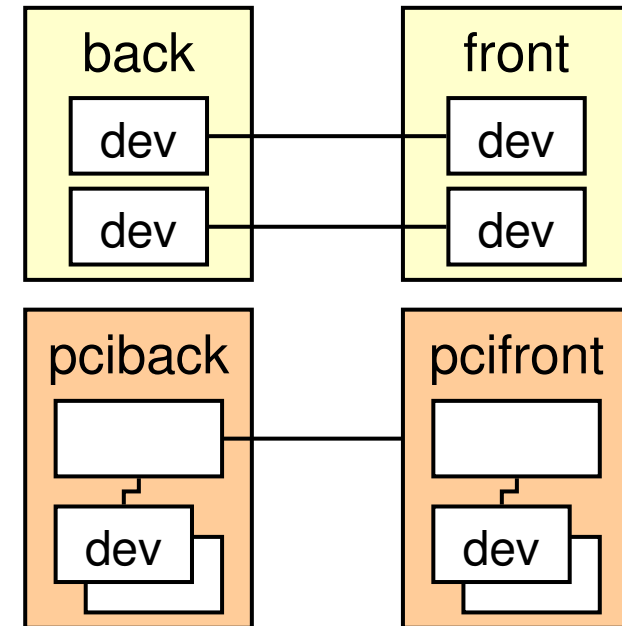
# Overview of PV PCI Pass-through (Driver Domain)



Xen Summit Tokyo 2008

# Approach

- Xenbus state transition
  - Normal PV split drivers
    - Each virtual device has dependent xenstore node and ring connection
    - Open connection on hot-add and close connection on hot-remove
  - pcifront / pciback
    - Single connection for guest virtual PCI platform
    - Each device is managed as attributes
- > Introduce sub-state for each device and reconfigure drivers while keeping connected



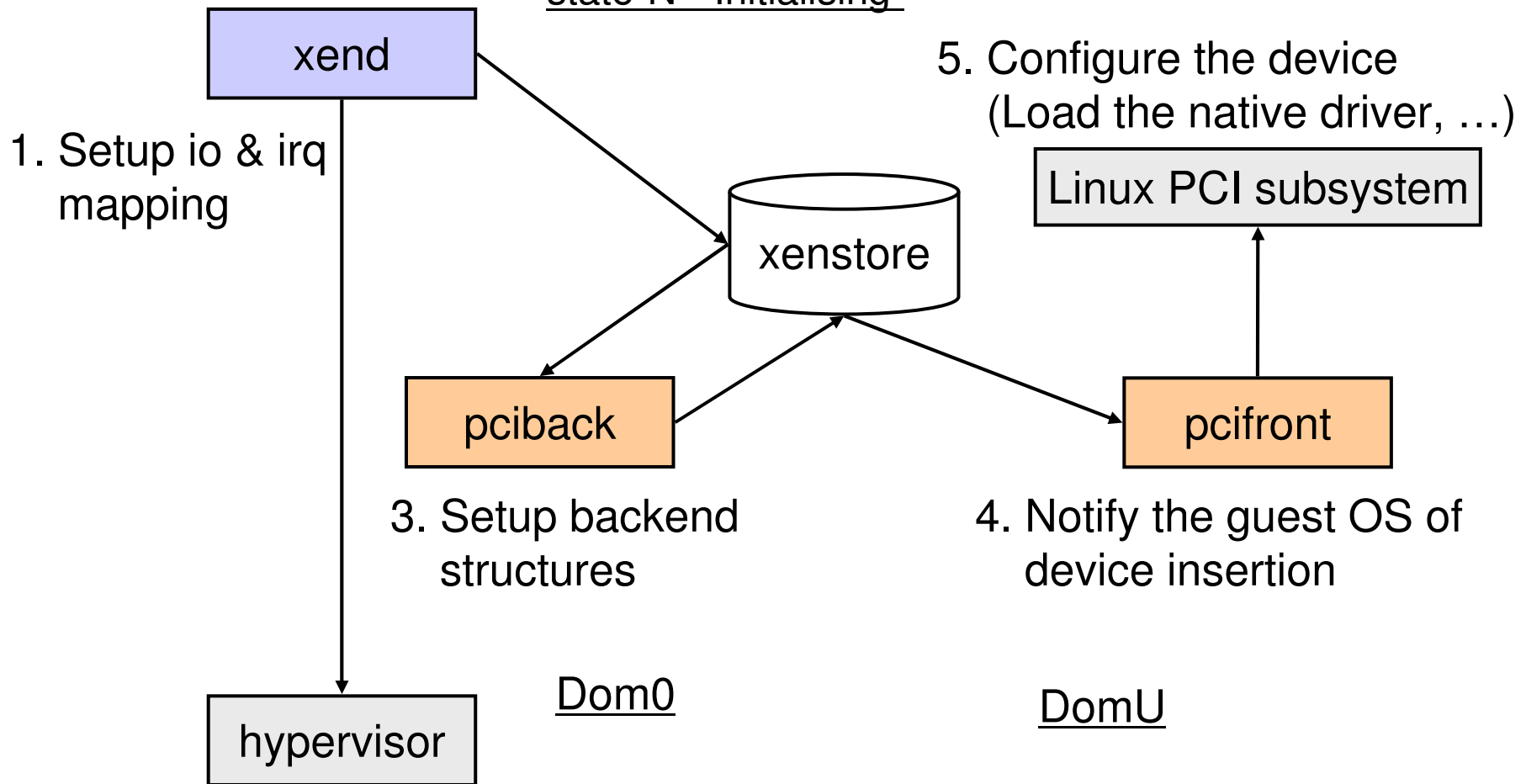
```
backend/pci/domid/0/  
  state -> global state  
  ...  
  dev-N -> phys B:D:F  
  vdev-N -> virt B:D:F  
  state-N -> sub-state  
  ...
```

Xen Summit Tokyo 2008

# Hot-Add Sequence

2. Indicate the device to attach and start reconfiguration

state-N="Initialising"



Xen Summit Tokyo 2008

# Hot-Remove Sequence

1. Indicate the device to detach and start reconfiguration

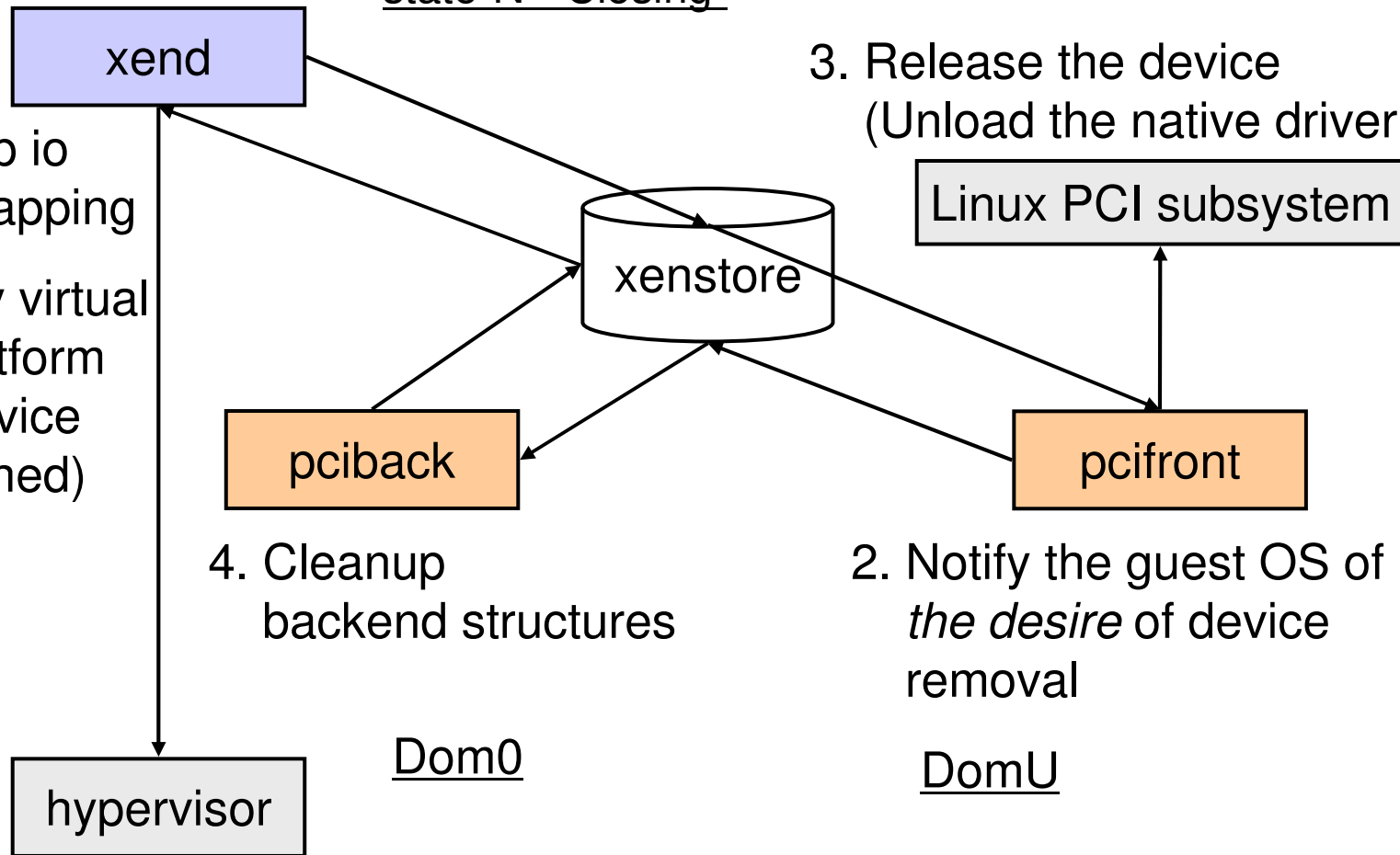
state-N="Closing"

3. Release the device  
(Unload the native driver, ...)

5. Cleanup io & irq mapping  
(6. Destroy virtual PCI platform if no device is attached)

4. Cleanup backend structures

2. Notify the guest OS of *the desire* of device removal



Dom0

DomU

Xen Summit Tokyo 2008

# Current Status and Future Prospects

# Current Status and Future Prospects

- Items included in Xen-3.3
  - PCI hot-plug for PV domains
  - PCI hot-plug for HVM domains (developed by Intel)
    - CLI integrated (xm pci-attach, pci-detach, pci-list)
  - Xen management API support for PCI pass-through is also available
- Remaining Tasks
  - Support more devices
  - PCI hot-plug for HVM with stub domains

Xen Summit Tokyo 2008

# Summary

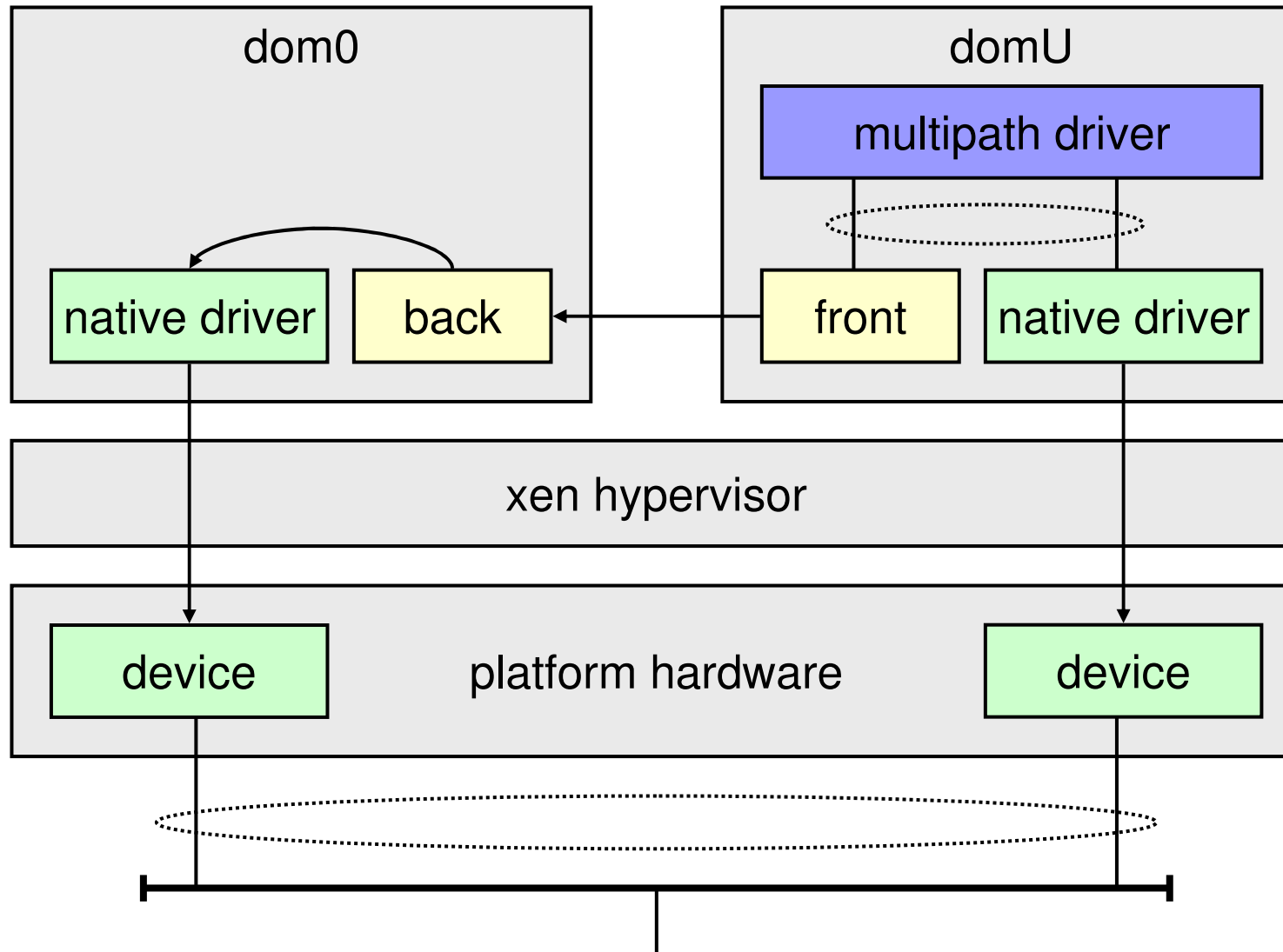
- PCI hot-plug for VM will add VM mobility and improve device utilization efficiency on PCI pass-through enabled virtualization systems
- Basic support of PCI hot-plug for PV domains has been developed and merged into Xen-3.3

Empowered by Innovation

**NEC**

# Backup Slides

# Bonding Physical Device and Virtual Device

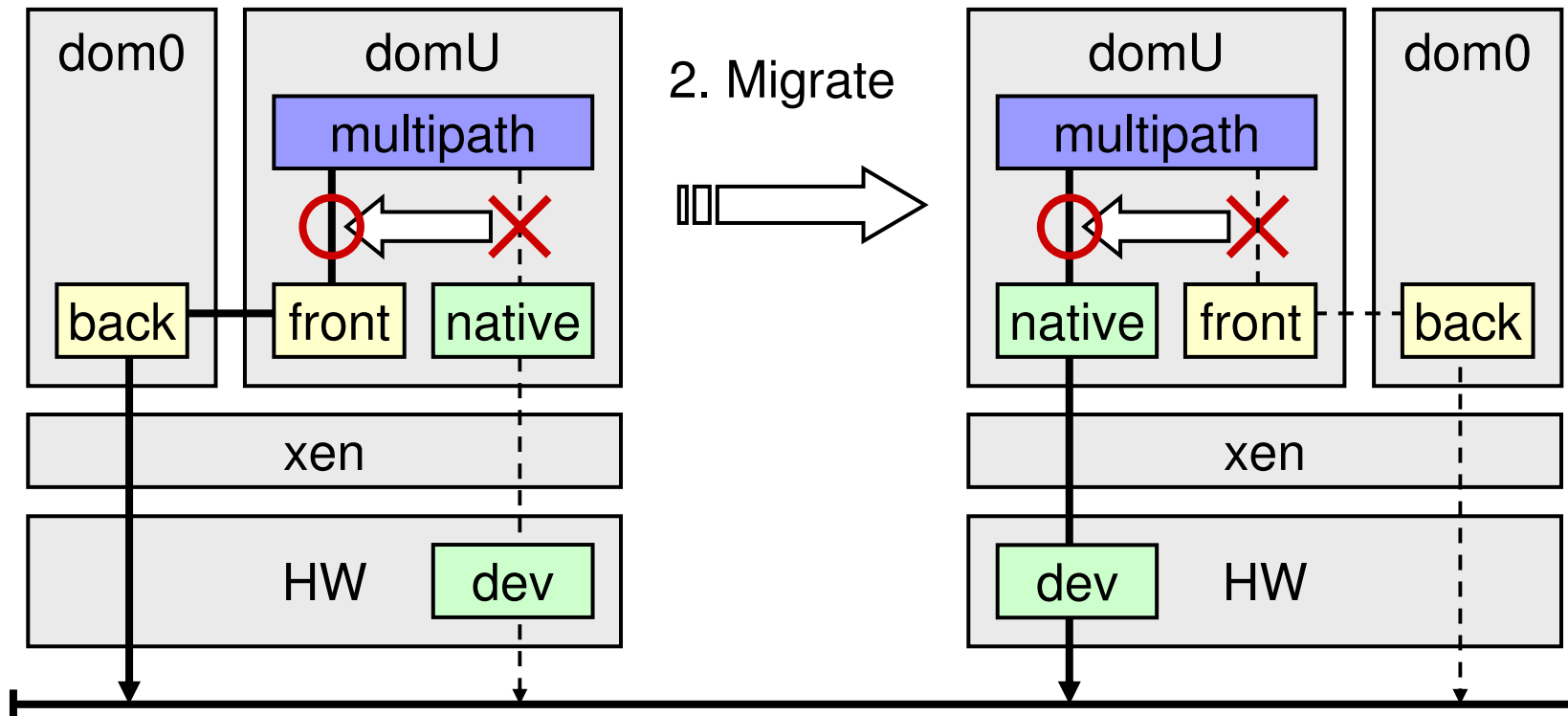


Xen Summit Tokyo 2008

# Migrating Device-assigned VM

1. Hot-remove  
the physical device  
and switch to  
the virtual path

3. Hot-add  
the physical device  
and switch to  
the physical path



Xen Summit Tokyo 2008