



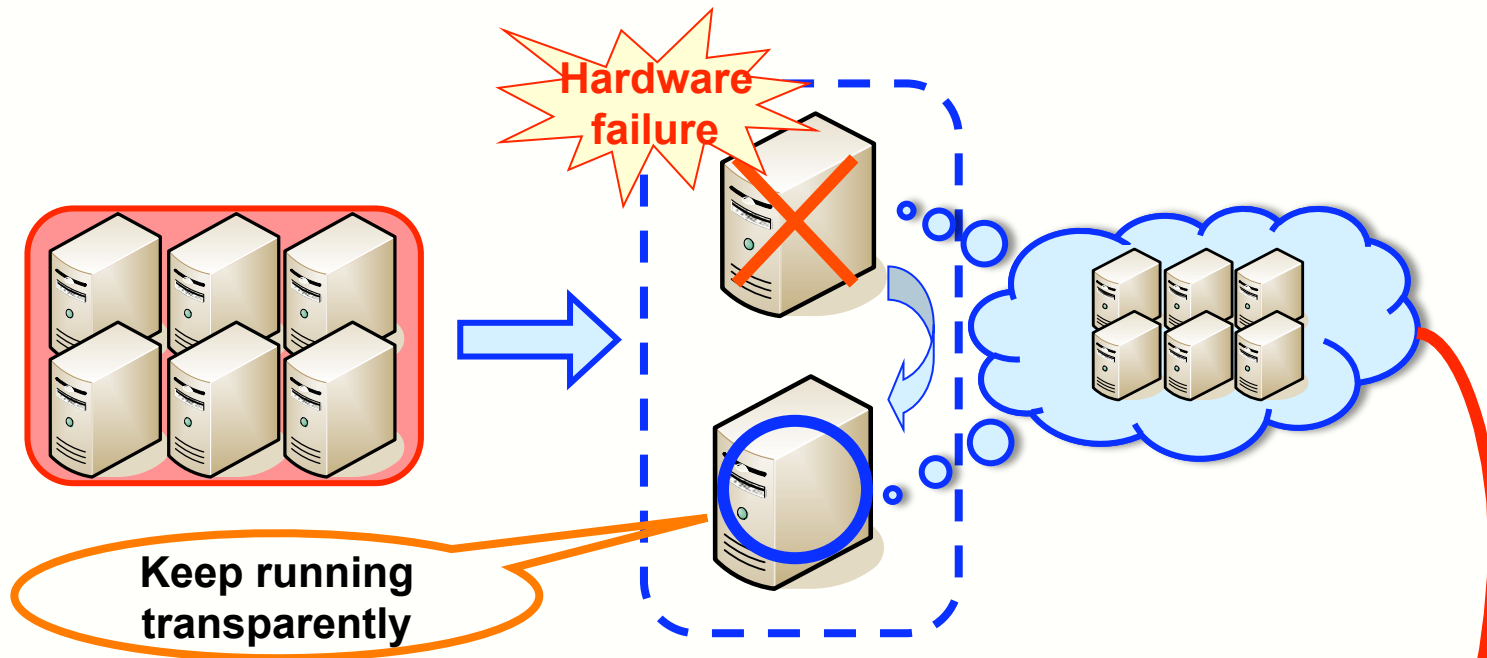
Kemari: Virtual Machine Synchronization for Fault Tolerance

Yoshi Tamura
NTT Cyber Space Labs.

2009/3/2

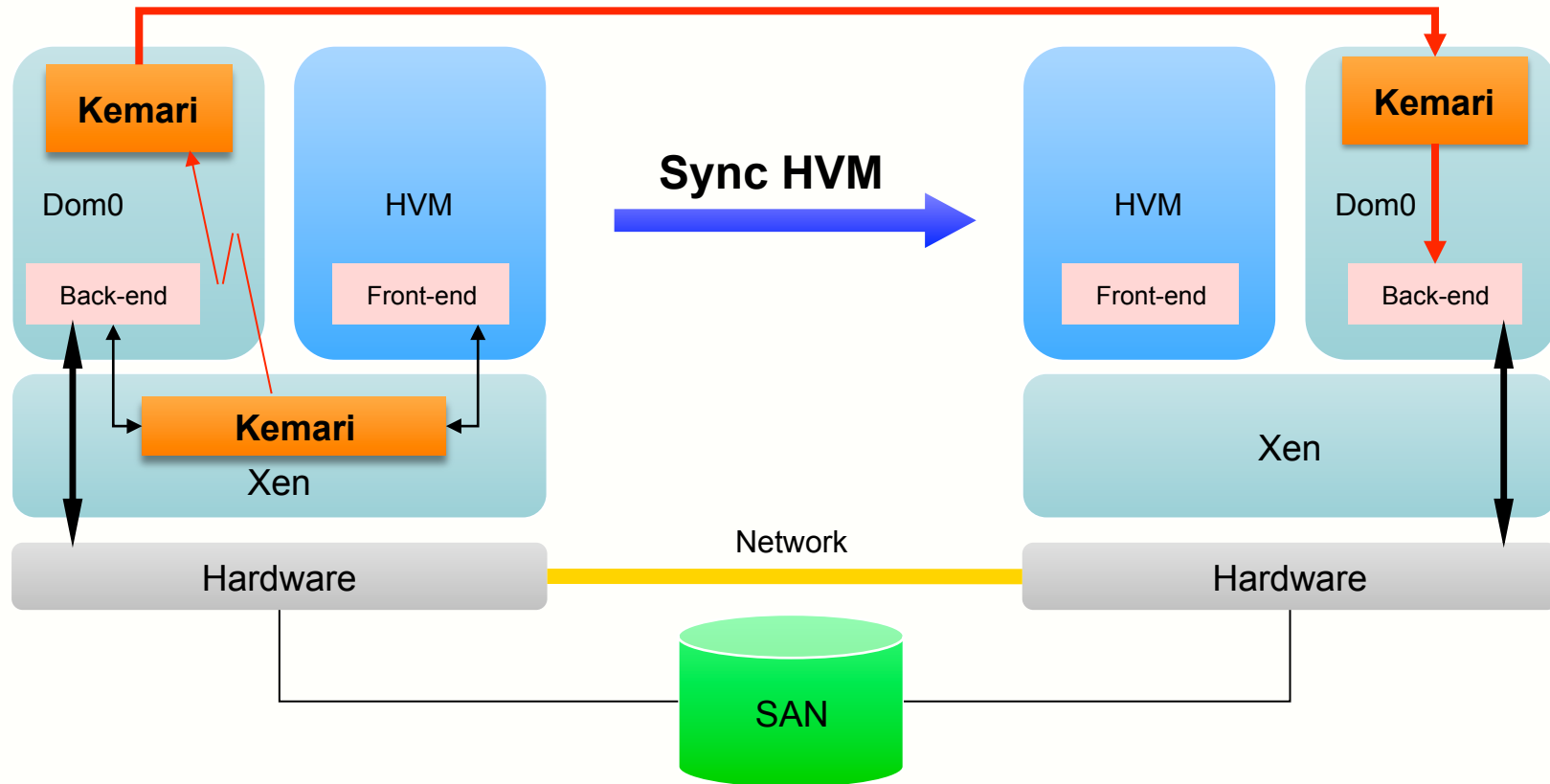
Our goal

Don't drop the ball! Don't drop the VMs!



Kemari: Virtual Machine Synchronization

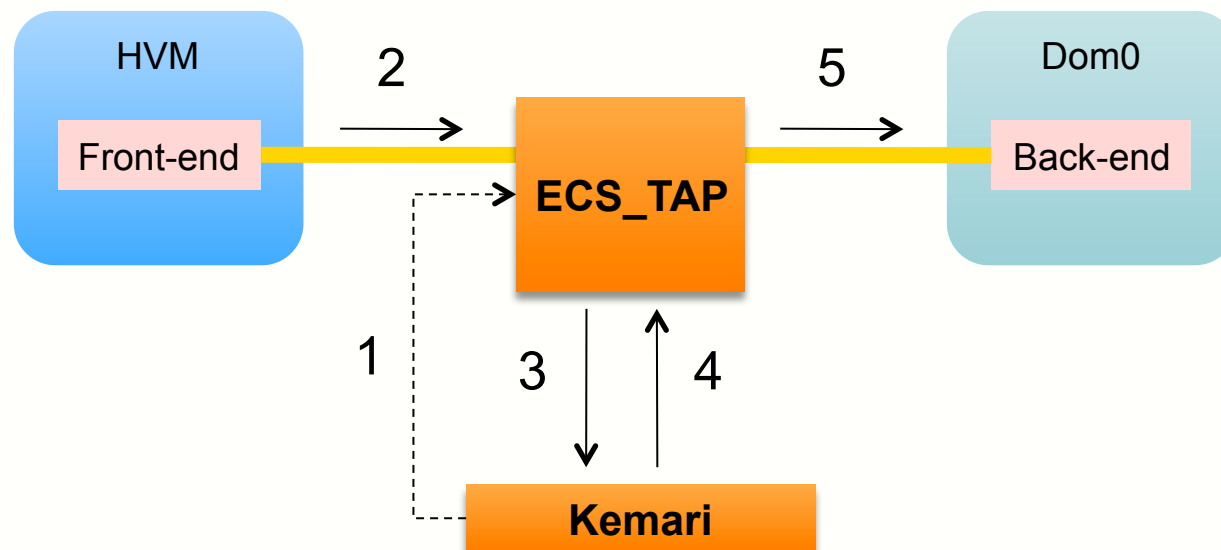
Architecture overview



- The core of the synchronization mechanism resides in hypervisor to synchronize virtual machines efficiently
- LOC \cong 4000 (hypervisor: 1000, Dom0+Tools: 3000)

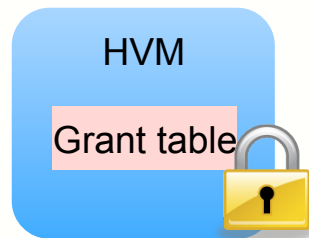
Event Channel tapping

- Simple but the key component of Kemari
- Monitors IN/OUT or Both
- Registered function is called on specific events
- Dynamically attachable
 - ▶ May be useful for measurements

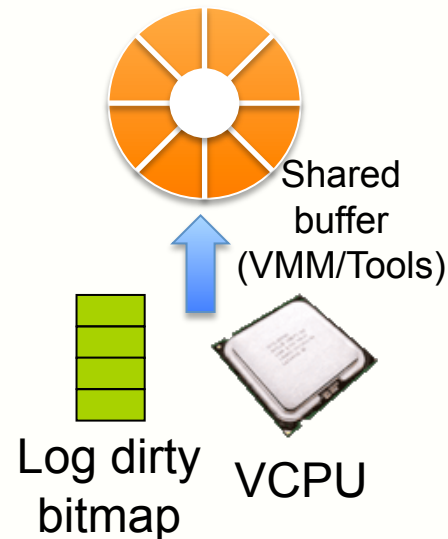


Transferring HVM

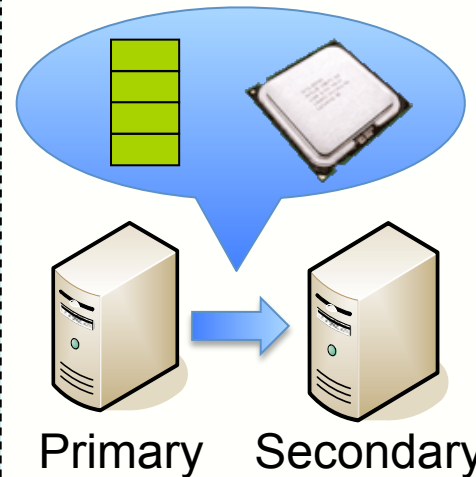
1 (VMM)



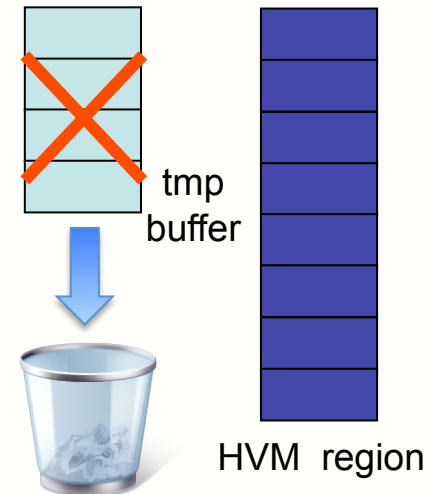
2 (VMM)



3 (Tools)



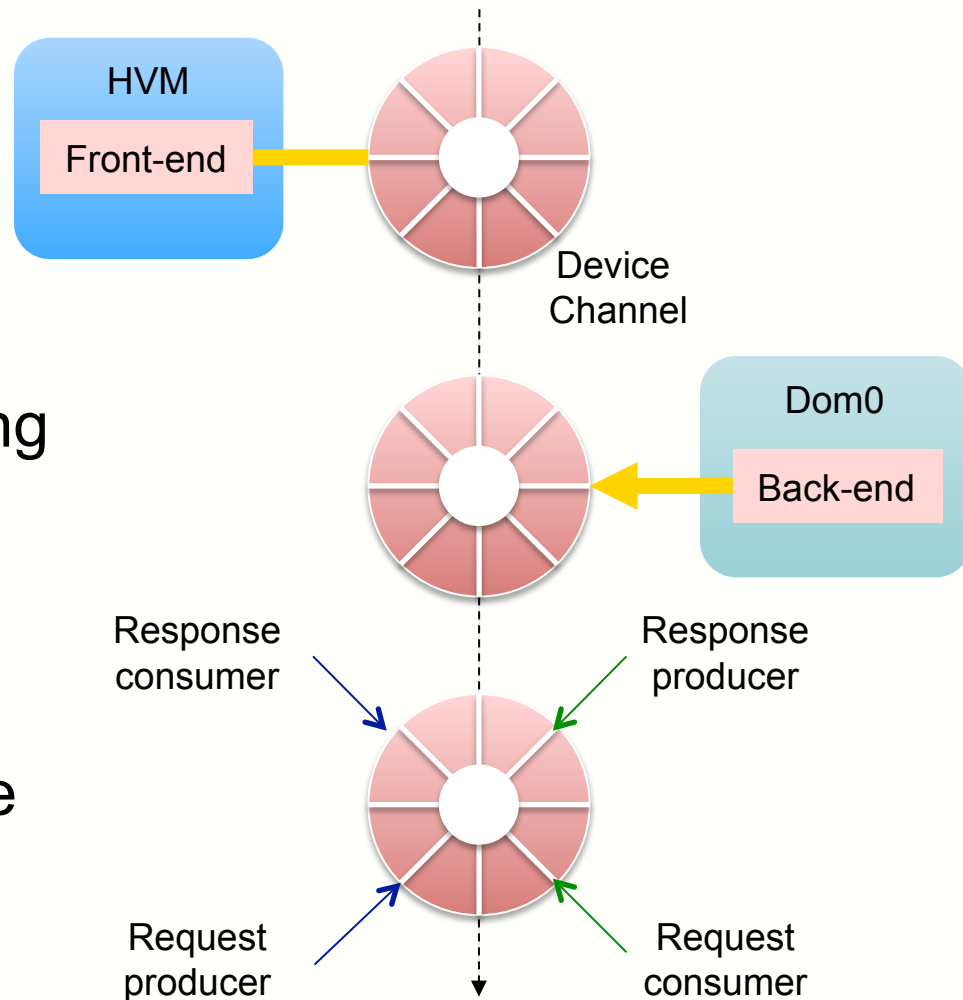
4 (Tools)



1. Pauses HVM and locks the grant tables. **No need to suspend!**
 - Grant tables are mapped at the last 4 pages of HVM region
2. Extracts dirtied pfn from the bitmap, copies pfns and the vcpu to the shared buffer, and notifies Tools via event channel
3. Maps dirtied pages, transfers pages and vcpu to the secondary
4. Secondary prepares temp buffers to rollback when failure is detected during transfer

Restoring para-virtualized devices

1. Device Channel is stored in HVM region



2. Attach the Back-end to the Device Channel using *BACK_RING_ATTACH* macro

3. Adjust producer and consumer indexes of the Back-end appropriately