Data Center Management

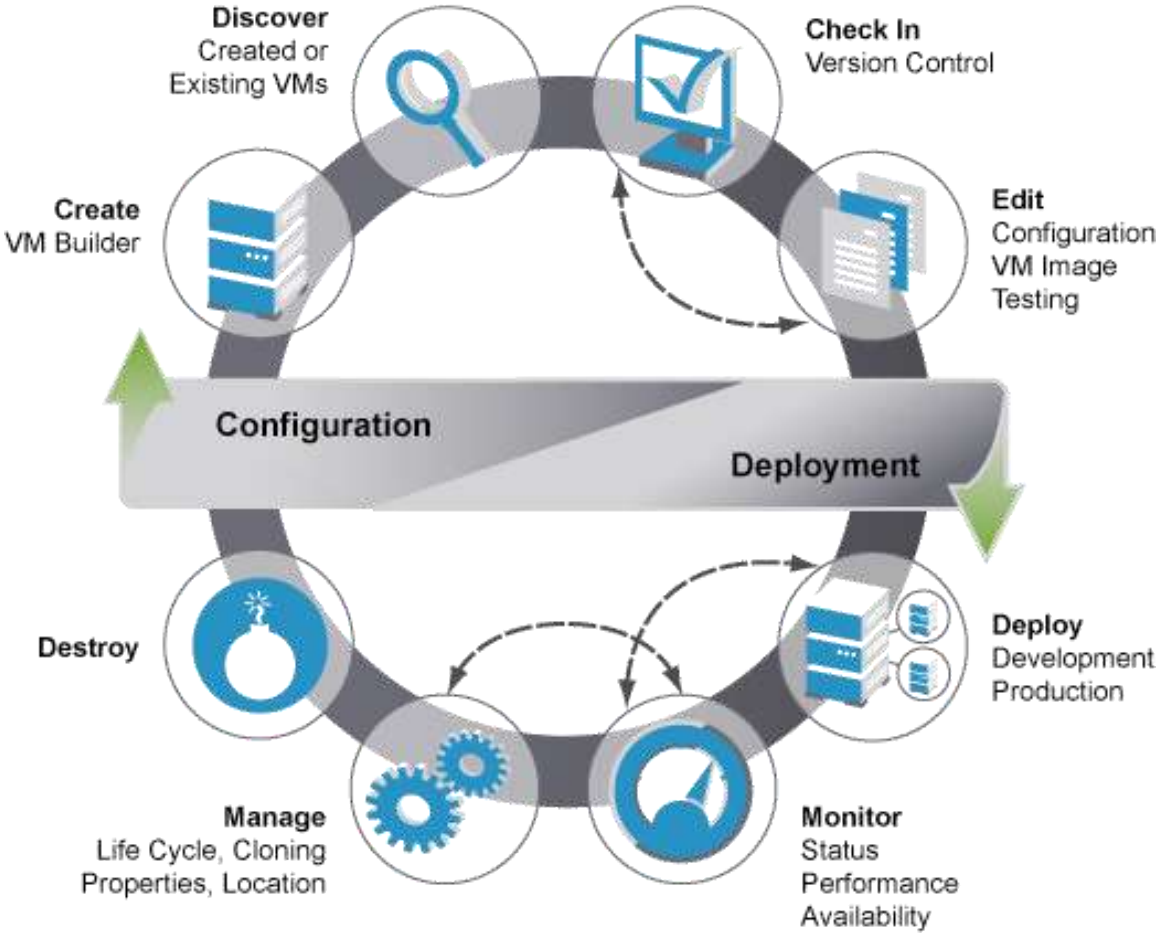
Cloud Computing

Outline

- Data center overview
- Virtual machine management
  - Consolidation
  - Migration
- Storage system
- Network system
- Dynamic provisioning
- Power management

VM MANAGEMENT

VM life cycle



VM consolidation

- Increase system utilization
  - VM consolidation is the way to share power/hardware/…
  - Hardware is getting cheaper, electricity is not
- Main challenge: memory system
  - Memory reclamation
  - Memory sharing
  - Memory compression

Memory overcommit

- Suppose each VM requests 4G RAM
  - But it usually uses 2.5 G RAM
- Suppose host machine has 48 G RAM
  - It can only support 12 VMs (48/4=12)
  - With memory overcommit
    it can suppose 19 VMs
  - The memory saved is 19*4-48 = 24G

Memory overcommit

- Memory is a non-renewable resource
- Secondary storage is really slow.
  - Many millions of CPU cycles in one disk seek
  - Process in guest accesses non-resident memory
- Memory reclaim
  - Which VM/process/page to reclamation?
- Memory sharing
  - Which pages can be shared?

Memory reclamation

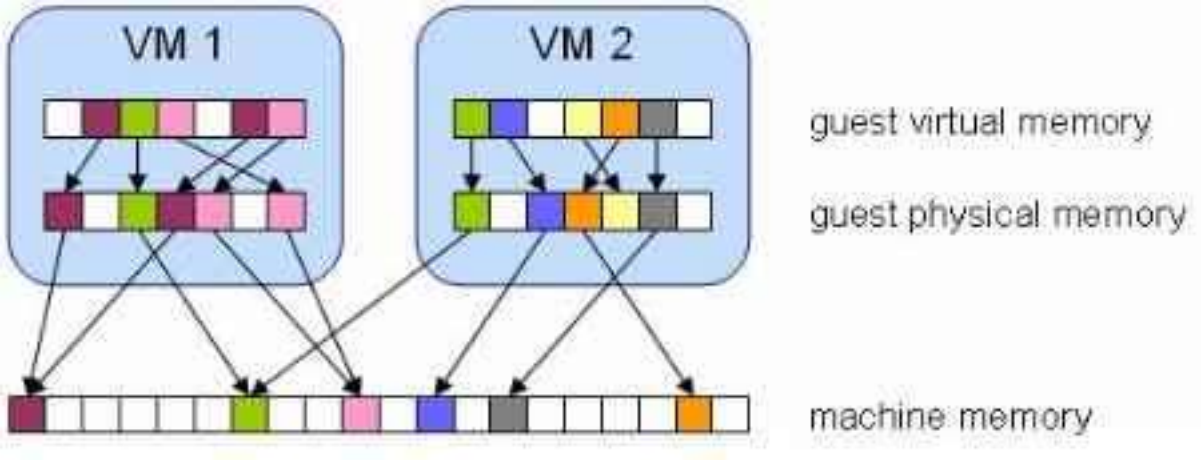
- Problem: when the physical memory is not enough, host machine (VMM) needs to reclaim memory from VMs
  - Which VM to reclaim?
  - Which pages to reclaim?
  - Double paging problem
- Solutions
  - Asynchronized page fault
  - Ballooning technique

Asynchronous page fault

- Host memory overcommit may cause guest memory to be swapped.
- When guest vCPUs access memory swapped out by a host its execution is suspended until memory is swapped back.
- Asynchronous page fault is a way to try and use guest vCPU more efficiently by allowing it to execute other tasks while page is brought back into memory.

Transparent page sharing

- **Multiple VMs share the same page**
  - **The content of the page are the same to all the VMs**
  - **Many system (OS) memory are the same**



Ballooning

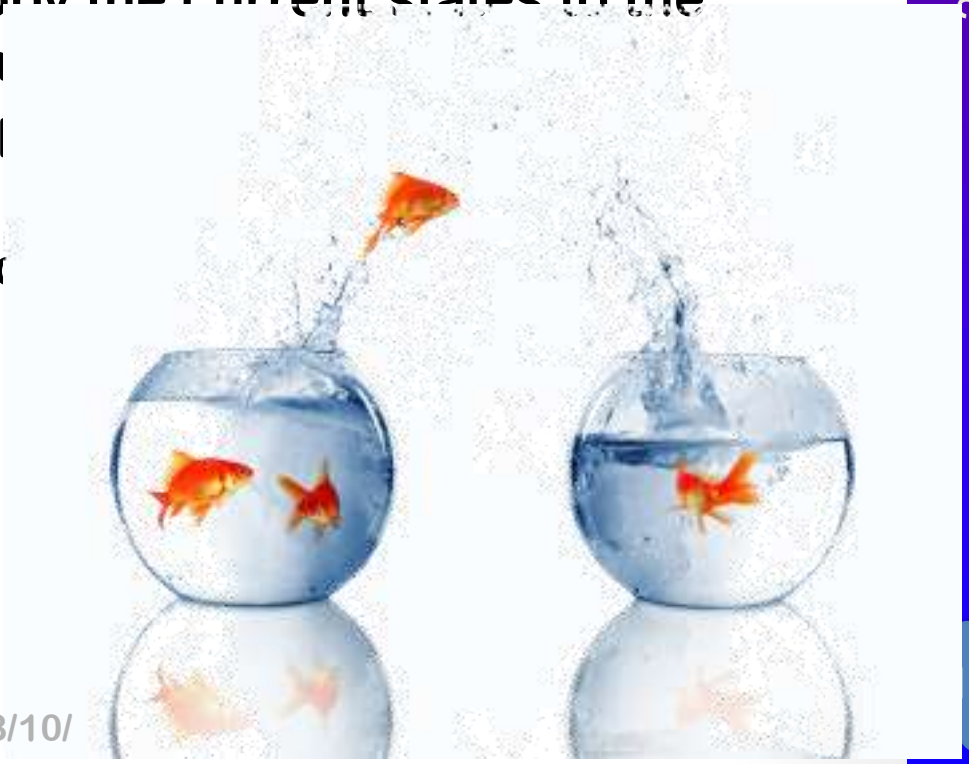
- Ballooning v A balloon is a pseudo-device driver to guest OS v Memory reclamation steps
- When host needs to reclaim memory from the guest OS, it inflates the balloon.
- Balloon requests memory in guest OS
- Guest OS needs to page out guest memory to satisfy the request.
- The requested memory by the balloon is reclaimed by the host.

Decide identical pages

- If there are N pages per VM and there are k VMs, the number of comparison will be $O(N^2k^2)$.
- Memory comparison of two pages is expensive.
  - Binary string comparison of two size m data, where m is the page size.
- Using hashing
  - Pages with the same hashcode may be identical.
  - Pages with different hashcodes must be different.
  - Reduce the unnecessary page comparisons.

What is VM migration?

- Moving a VM from one physical machine to another one
  - Cold migration: VM stops to execute any currently working program and copy the current states to the machine where the VM m
  - Live migration: During t migration process, the execution in VM might g on without stopping the execution of programs.



http://criterionglobal.wordpress.com/2008/10/

Why VM migration?

- Consolidate resources
- Load balance
- System maintenance
- Performance improvement
- User's carte blanche of VM
- Why live migration?
  -
  -

What to migrate?

- The VM status needs be moved to new physical machine
  - Register, memory, hard disk, …
  - Data in cache are usually flushed before migration
  - Data in hard disk may not be attached to the physical machine
- VM managing information.
  - Virtualized hardware info.

Problems of migration

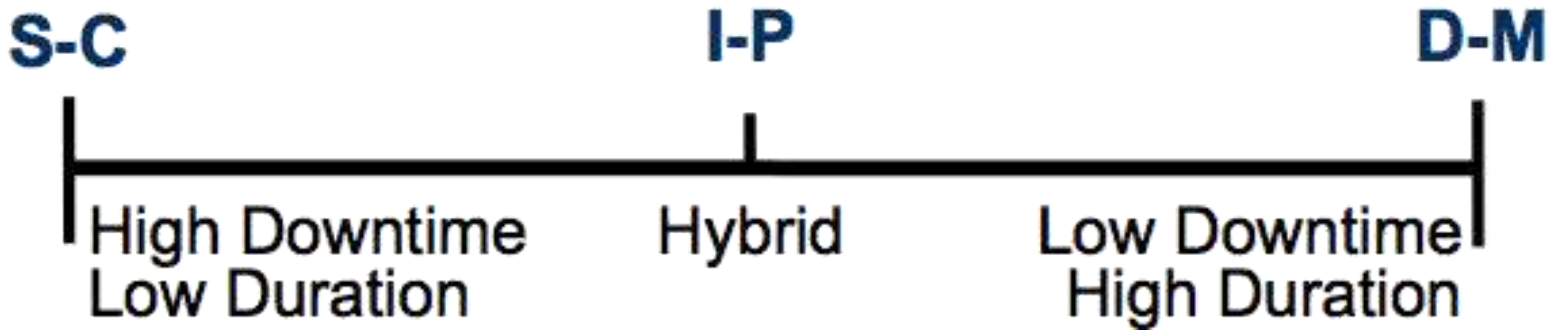
- Performance of server migration
  - The system state includes memory, cache, register,…
  - The memory size can be very large
  - Minimize the downtime
- The data/storage migration
  - Block device migration
  - NAS(Network-attached storage) redirection
- The network migration
  - Maintain the network connection
  - Maintain the LAN structure

Two major concerns

1. Downtime: the period during which the service is unavailable due to there being no currently executing instance of the VM
   - This period will be directly visible to clients of the VM as service interruption.
2. Total migration time: the duration between when migration is initiated and when the original VM may be finally discarded.
   - The source host may potentially be taken down for maintenance, upgrade or repair.

Migration methods

- Stop-and-copy (S-C)
- Demand-migration (D-M)
- Iterative precopy (I-P)



Stop and copy

- Procedure
  1. Stop source VM
  2. Copy all pages over the network
  3. Start destination VM
- Longest service downtime
- Shortest migration duration

Demand migration

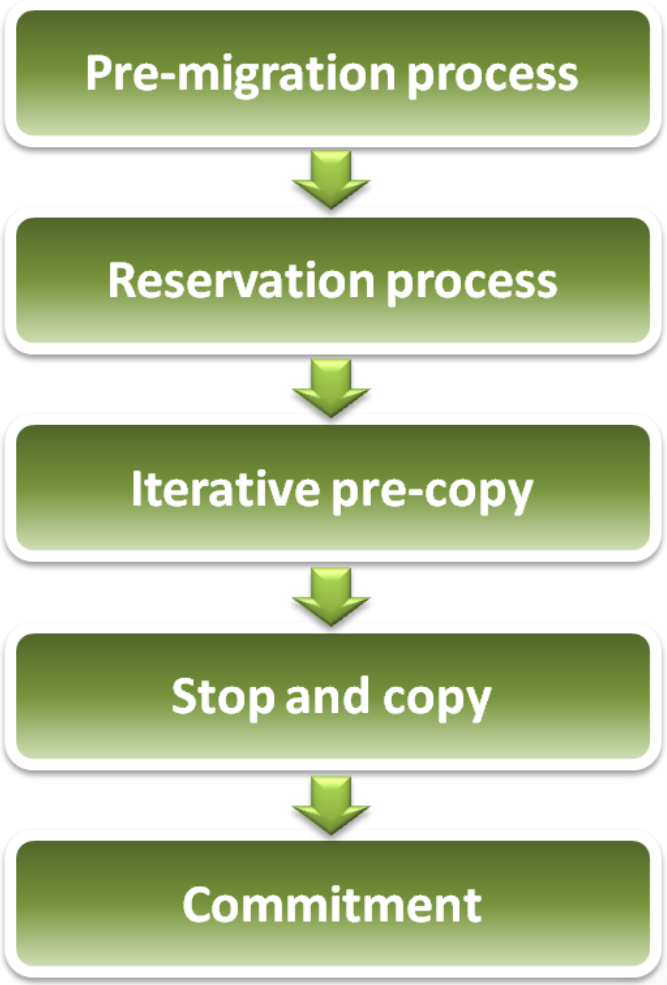
- Procedure
  1. Copy over critical OS structures
  2. Start destination VM
  3. Page faults trigger network copy
- Shortest Service Downtime
- Longest Migration Duration

Iterative precopy

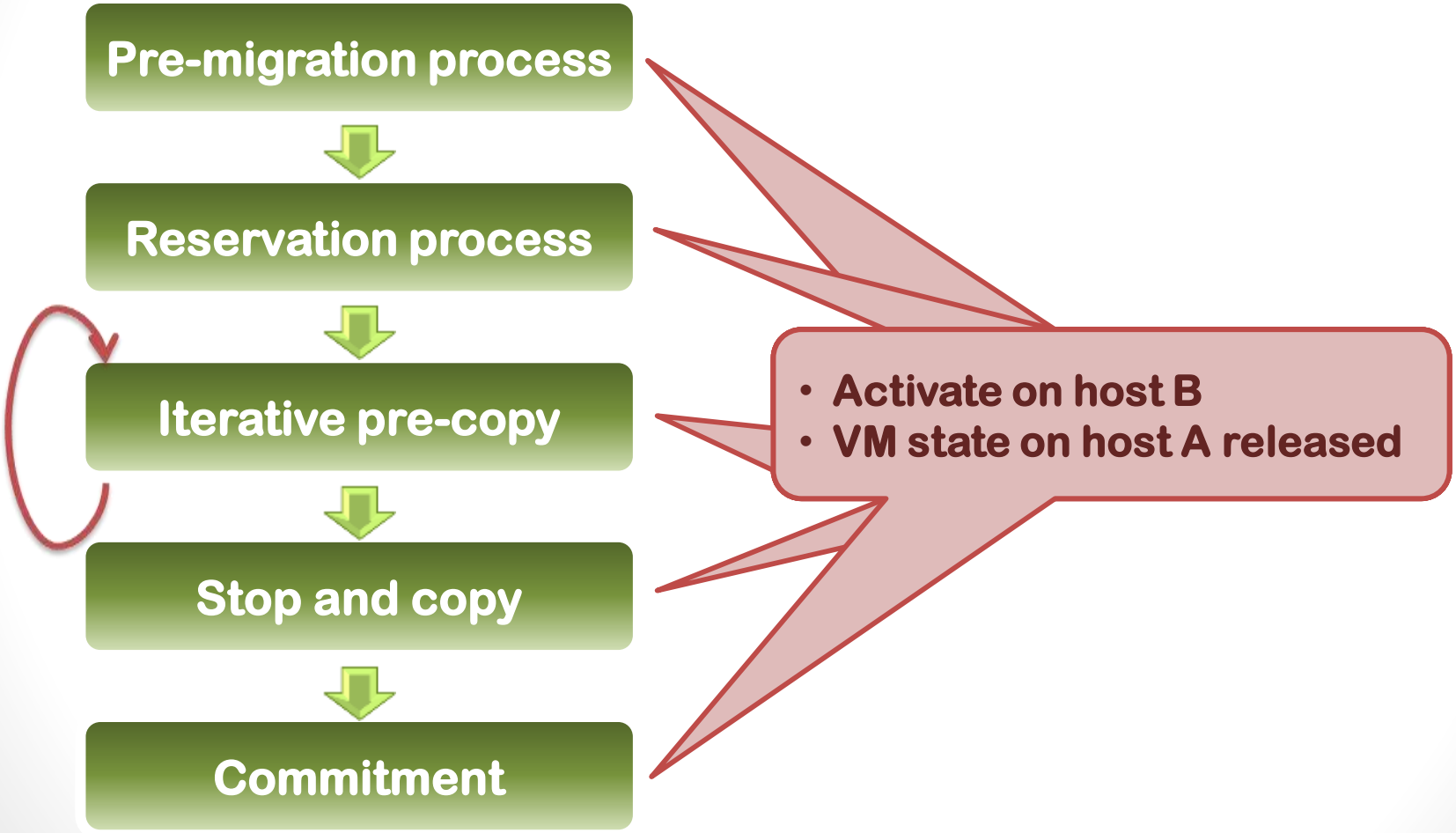
- Procedure
  1. Iteratively copy pages over network
  2. Keep copying dirtied pages until threshold
  3. At threshold, stop source VM, copy remaining pages, start destination VM
- Balances service downtime and migration duration
- Method used by VMware/Xen

Live migration

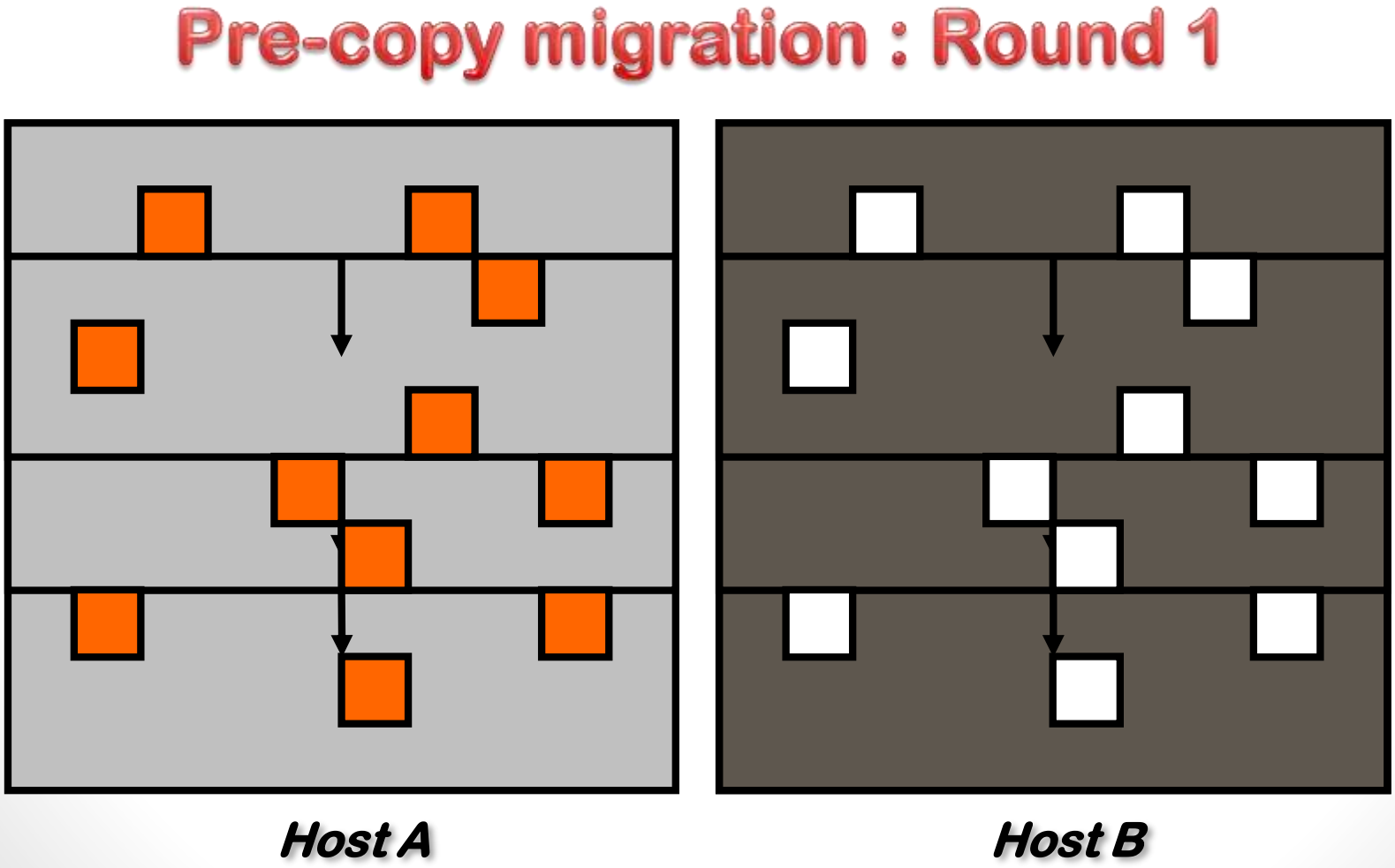
- Relocation strategy :
  1. Pre-migration process
  2. Reservation process
  3. Iterative pre-copy
  4. Stop and copy
  5. Commitment



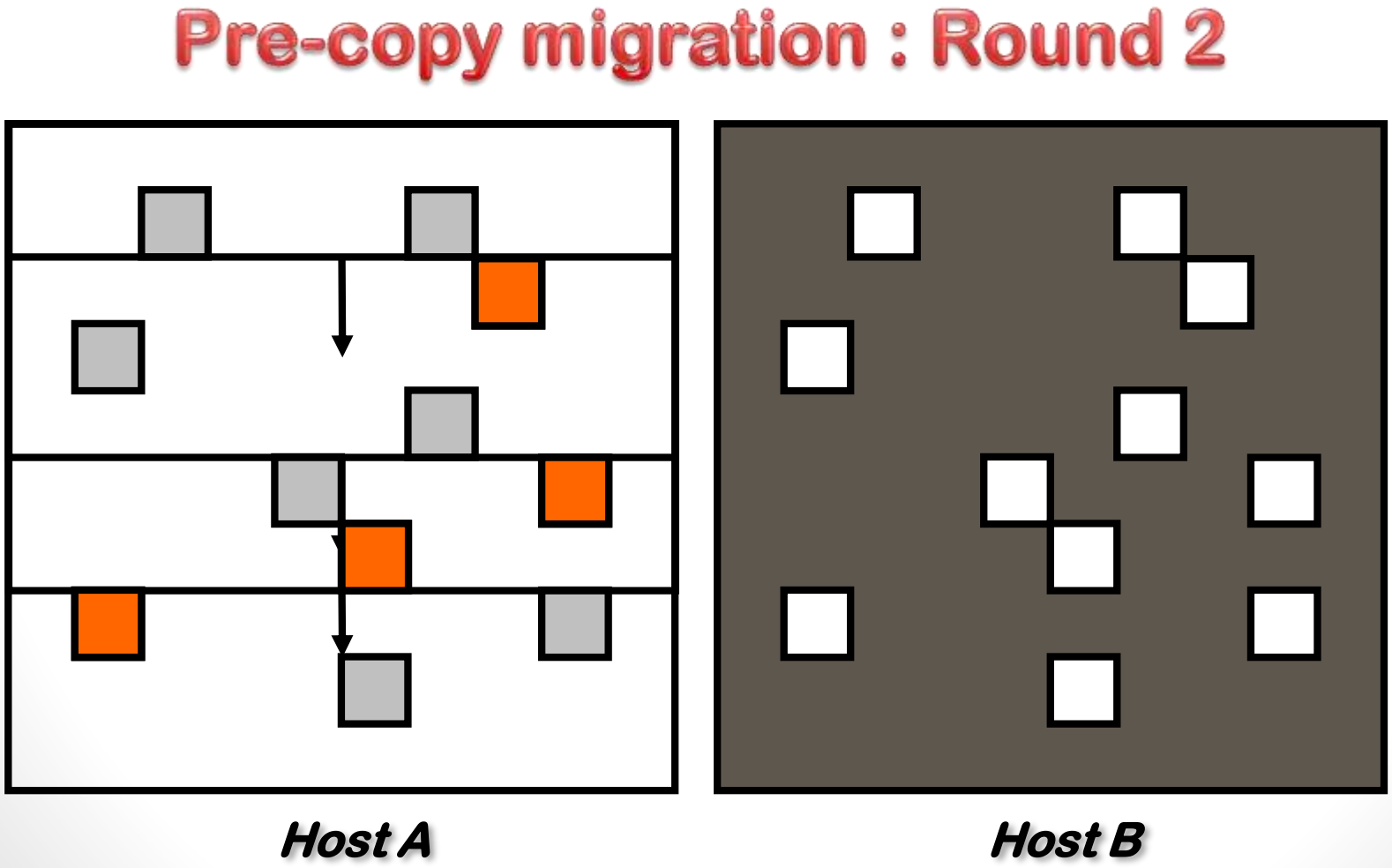
Live Migration Technique



Example: Live migration

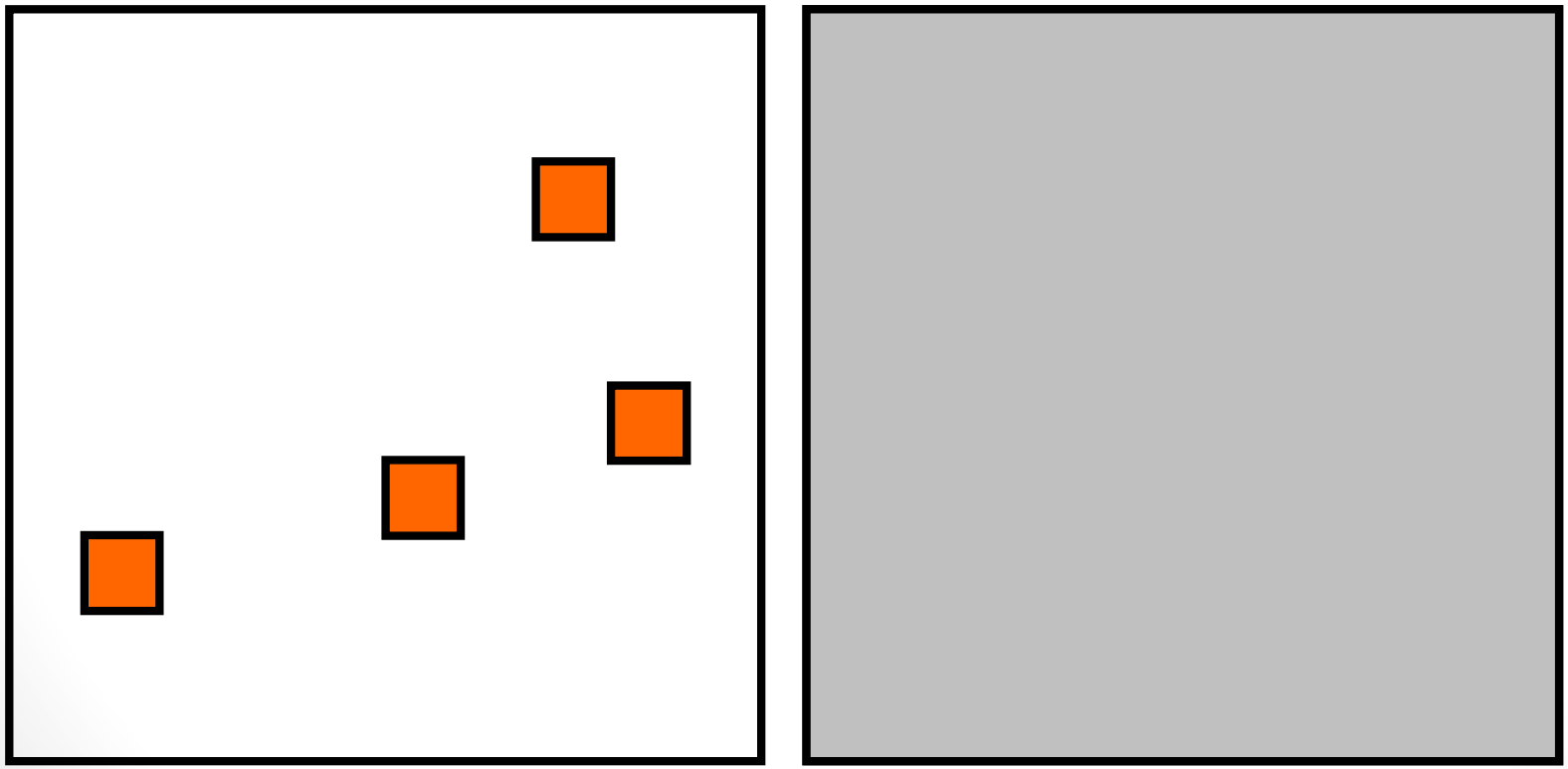


Example: Live migration



Example: Live migration

Stop and copy : Final Round



Problems

- Migration of block devices
  - Some harddisk is still attached on host
  - Massive data to migrate
- Network forwarding/redirecting/tunneling
  - Ongoing network transmissions.
  - Migration over WAN*
- Hardware devices
  - Some data in hardware buffer need be migrated too.
  - Hardware assistant virtualization.