



BITS Pilani

Cloud Computing

Session 11-12

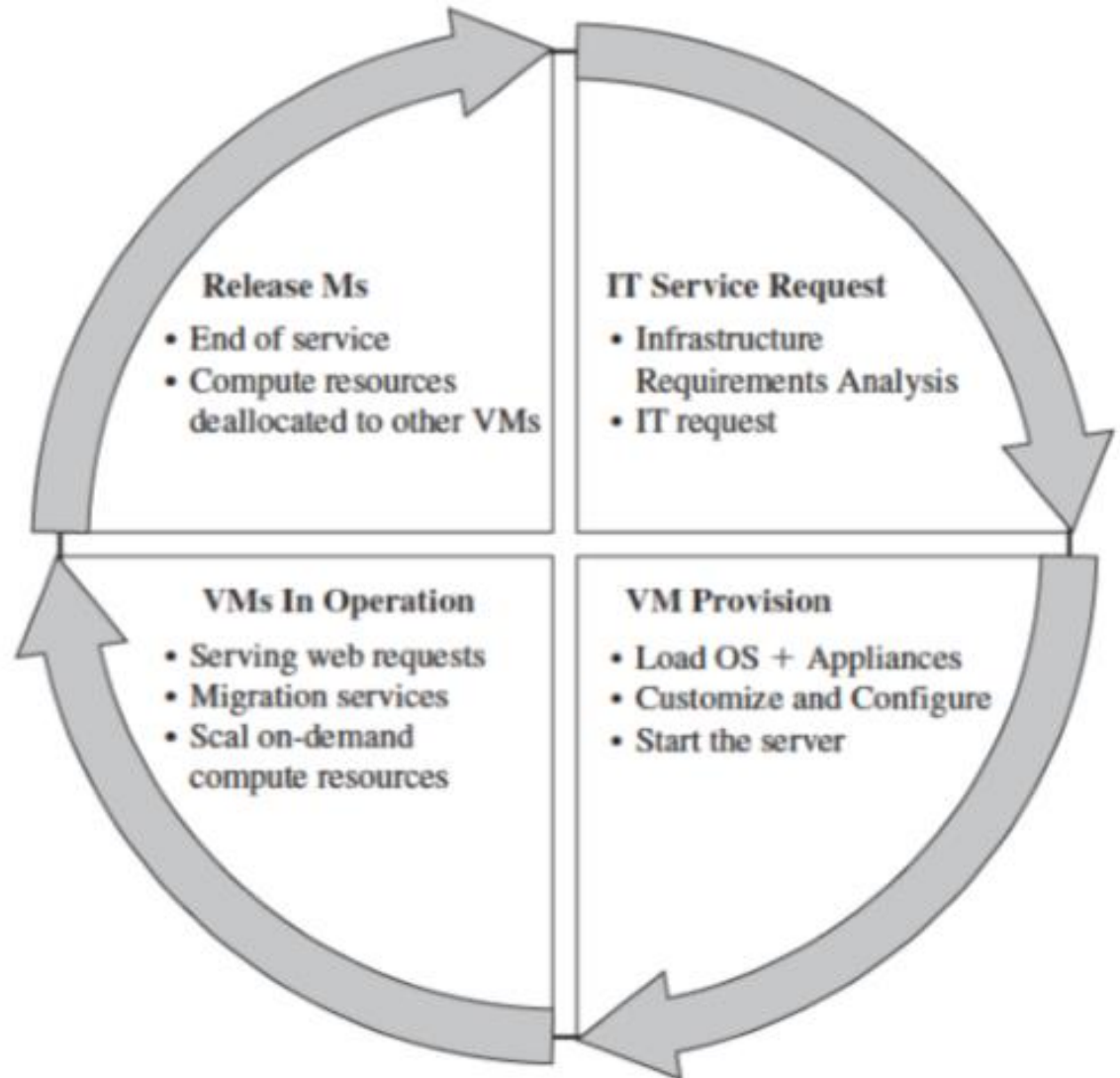
VM Provisioning & Migration

Agenda

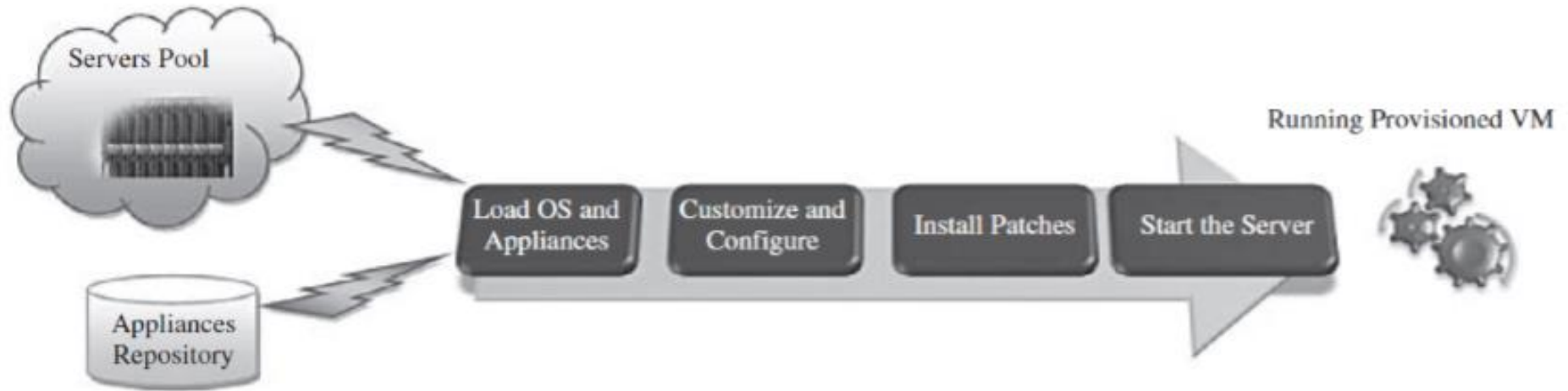
- VM Provisioning
- VM Migrations
 - Cold / Regular
 - Hot / Live
 - Flavors

Virtual Machine Provisioning Life Cycle

1. IT Service Request
2. VM Provision
3. VM in Operation
4. VM Release



VM Provisioning Process



1. Select a server from a pool of available servers along with the appropriate OS template
2. Load the appropriate software.
3. Customize and configure the machine (e.g., IP address, Gateway) to an associated network and storage resources.
4. Finally, the virtual server is ready to start with its newly loaded S/W.

VM Provisioning

- Manually installing an OS,
- Using a preconfigured VM template,
- Cloning an existing VM
- Importing a physical server or a Server from another hosting platform.
- Physical servers can also be virtualized and provisioned using P2V (Physical to Virtual)

VM Provisioning using templates

- Using template reduces the time required to create a new virtual machine.
- Administrators can create different templates for different purposes.

For example –

- Vagrant provision tool using VagrantFile (template file)
- Heat – Orchestration Tool of openstack (Heat template in YAML format)
- Instance creation in cloud, Load balancer in cloud

Enables the administrator to quickly provision a correctly configured virtual server on demand.

VM Migration

What is it ?

The process of moving a virtual machine from one host server or storage location to another.



Image courtesy: Britannica

VM Migration

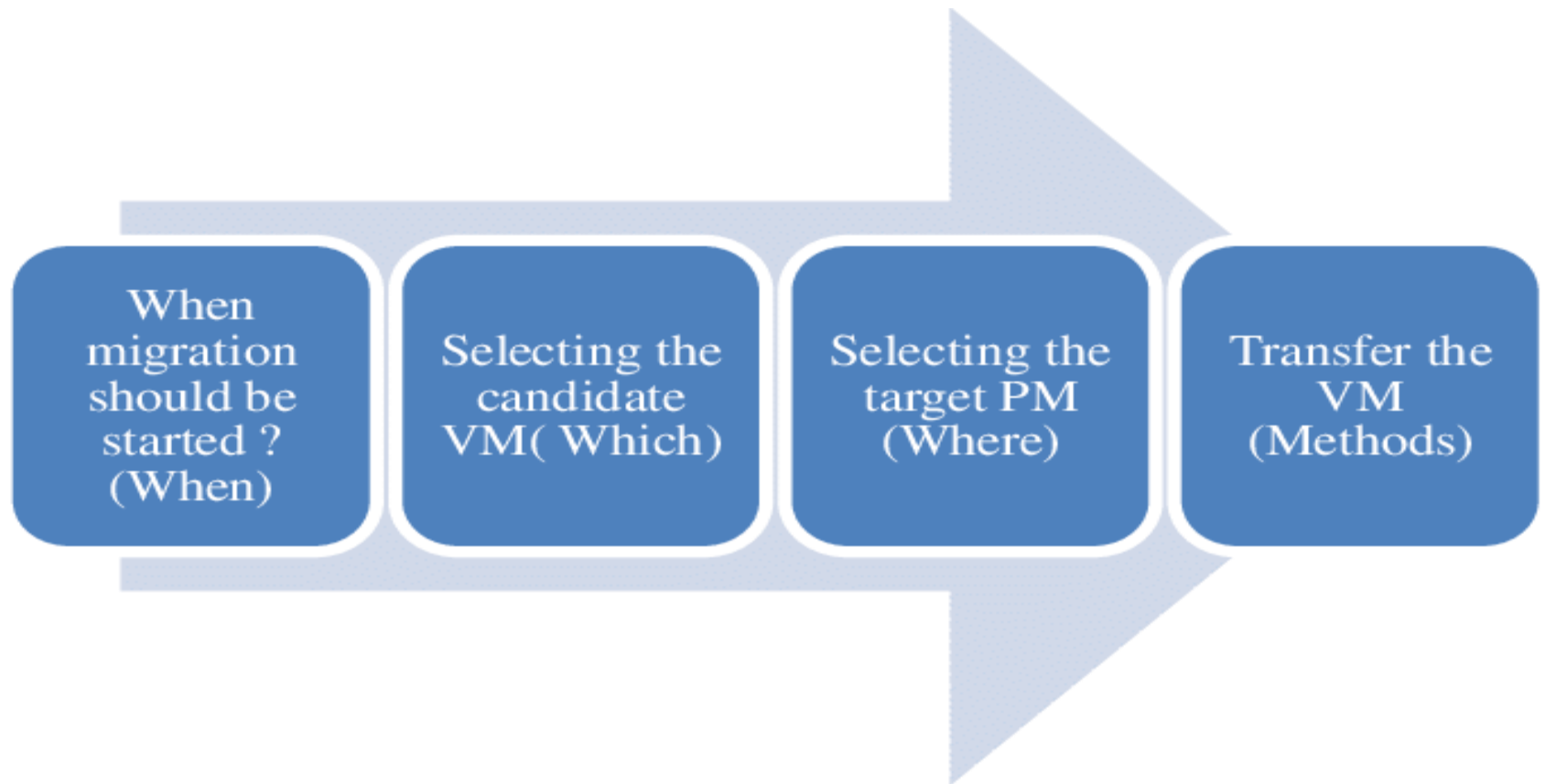


Image courtesy: researchgate.net

VM Migration Flavours

Main Classification

1. Hot/live migration
2. Cold/regular migration



Facilitating VM Migration

All key machines' components, such as CPU, storage disks, networking, and memory, are completely virtualized, thereby facilitating the entire state of a virtual machine to be captured by a set of easily moved data files.

Why Migration

- **Load Balancing**

- Move VMs from highly loaded servers to lightly loaded servers

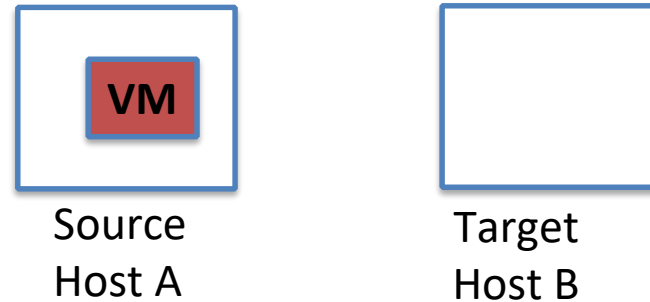
- **Server maintenance**

- When server needs to be upgraded

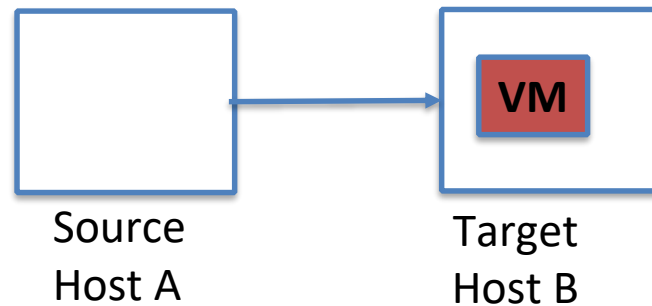
- **Energy savings**

- Move out VMs before shutting down servers to reduce energy usage
-

Migrating a VM from Host A to Host B



Before Migrating a VM from host A to host B



After Migration of VM from host A to host B

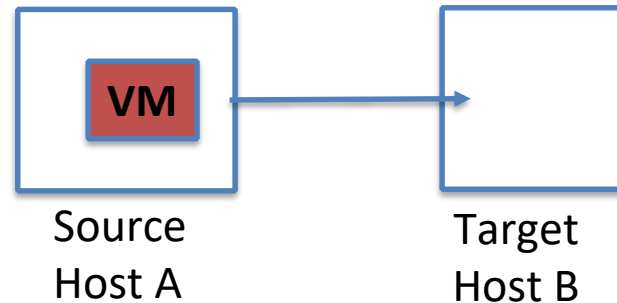
Cold /Regular Migration

- The migration of a powered-off virtual machine
 - Done in the following steps:
 - 1a. If the option to move to a different datastore was chosen,
 - the configuration files, including the NVRAM file (BIOS settings), and log files are moved from the source host to the destination host's associated storage area.
 - 1b. If the option to move the virtual machine's disks, these are also moved.
 - 2. The virtual machine is registered with the new host.
 - 3. After the migration is completed, the old version of the virtual machine is deleted from the source host.
-

Why Live Migration

- Live migration involves migrating a running/alive virtual machine from one host to the other
 - **Why live?**
 - To keep long-running jobs alive
 - To keep network connections alive
 - Broadly, to avoid disruptions to users of VM
 - No noticeable interruption to the VM (Ideally)
 - **Why VM?**
 - Why not migrate individual processes?
 - Process migration may leave residual dependencies (state) at source host
 - E.g. System call redirection, shared memory, open files, inter-process communication, etc.
-

Live Migration: High Level Steps



Migrating a VM from host A to host B

1. Setup target host B. Reserve resources for the VM
 2. **Push**: Push some memory of VM from A to B
 3. **Stop-and-copy**: Stop the VM at A, copy CPU context, and some memory
 4. **Pull**: Start VM at host B, pull any further memory required from A
 5. Clean up state from host A, migration complete
- **Total migration time**: Time for steps 2,3,4
 - **Service downtime**: Time for step 3

Slide text courtesy: IITB

Performance Goals in Live Migration

- Minimizing Downtime
- Reducing total migration time
- Avoiding interference with normal system activity
- Minimizing network activity

Evaluate metrics:

- Impact on application performance
 - Network bandwidth
 - Total pages transferred
-

Challenges in Live Migration

- VMs have lots of state in memory
- Some VMs have soft real-time requirements :
 - For example: web servers, databases and game servers, etc.
 - Need to minimize down-time

Types of Live Migration

1. Pure Stop and Copy
2. Pre-Copy
3. Post-Copy
4. Hybrid-approach of pre-copy and post-copy

Pure Stop-and-Copy



- Freeze VM at source
- Copy the VM's memory contents to target
- Restart VM at target
 - Long downtime
 - Minimal total migration time = downtime

Pre-copy Migration

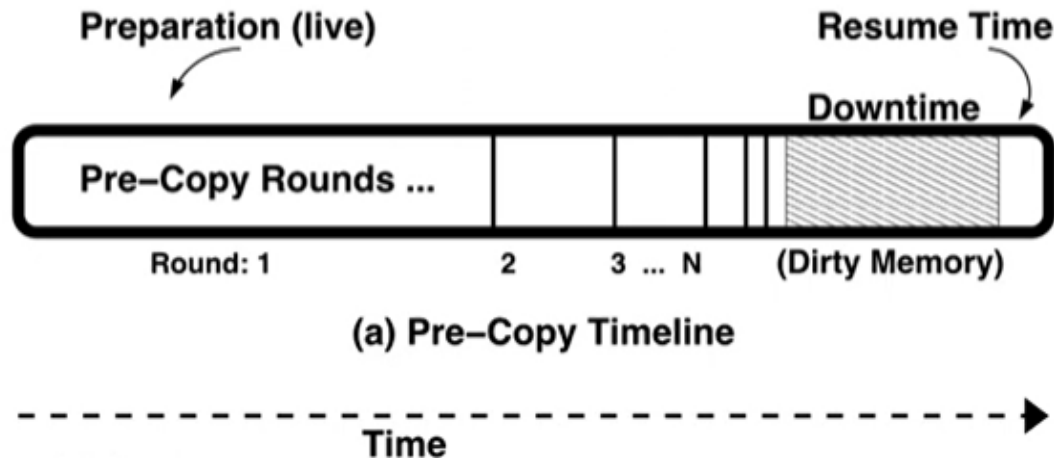


1. Pre-migration process
2. Reservation process
3. Iterative copy
4. Stop and copy
5. Commitment

Pre-copy Migration



Pre-copy migration



DON'T freeze VM at source

Let the VM continue to run

Copy VM's pseudo-physical memory contents to target over multiple iterations

Pre-Copy Technique

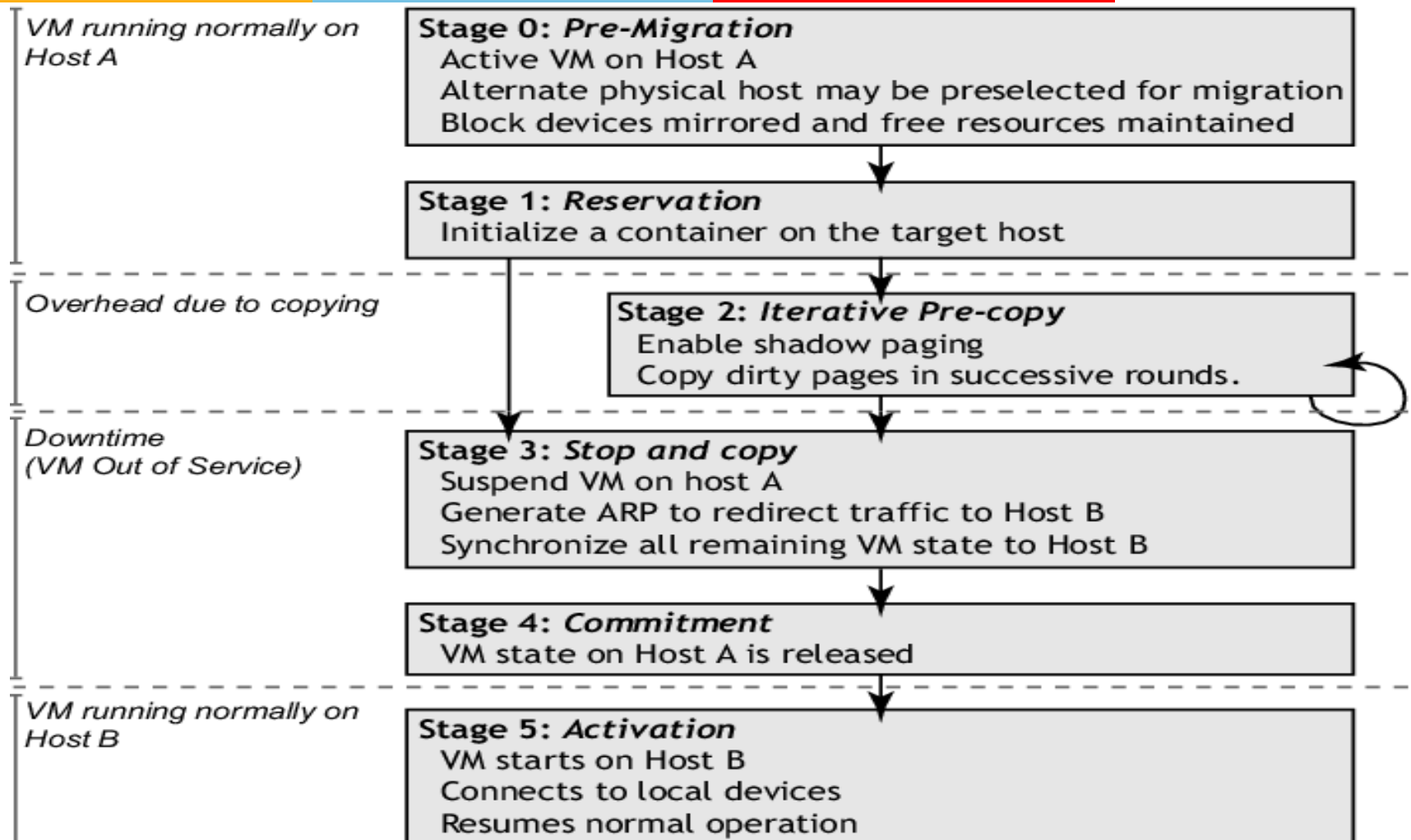


Image courtesy: "Live Migration of Virtual Machines", Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen†, Eric Jul†, Christian Limpach, Ian Pratt, Andrew Warfield

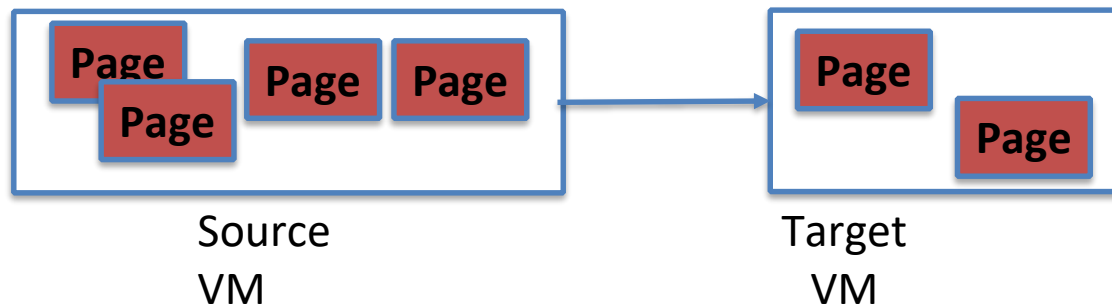
Pre-copy Migration



- First iteration → copy all pages.
- Each subsequent iteration → copy pages that were dirtied by the VM during the previous iteration
- Do a short stop-and-copy when number of dirty pages is "small enough".
- But what if number of dirty pages never converges to a small enough number?
 - After a fixed number of iterations, give up and stop-and-copy.

How do we track dirtied pages?

- Mark the VM's memory pages as read-only after each iteration.
- Trap write operations via hypervisor and track dirtied pages.
- Reset after each iteration
- Works well as long as writes are infrequent



Steps of Tracking Dirty Pages

Shadow page table constructed on demand for every round

- Dirty bitmap maintained for every round
- Any page access by guest, page fault to hypervisor, shadow page table updated
- PTE marked as read-only by default in shadow
- If valid write access, shadow PTE marked writeable, page marked dirty in bitmap
- At end of round, dirty pages are marked for transfer in control software
- Last set of dirty pages copied in stop-and-copy
- Shadow page table and dirty bitmap reinitialized after every round
- Guest page table in target host changed based on new physical addresses

Pre-Copy Performance

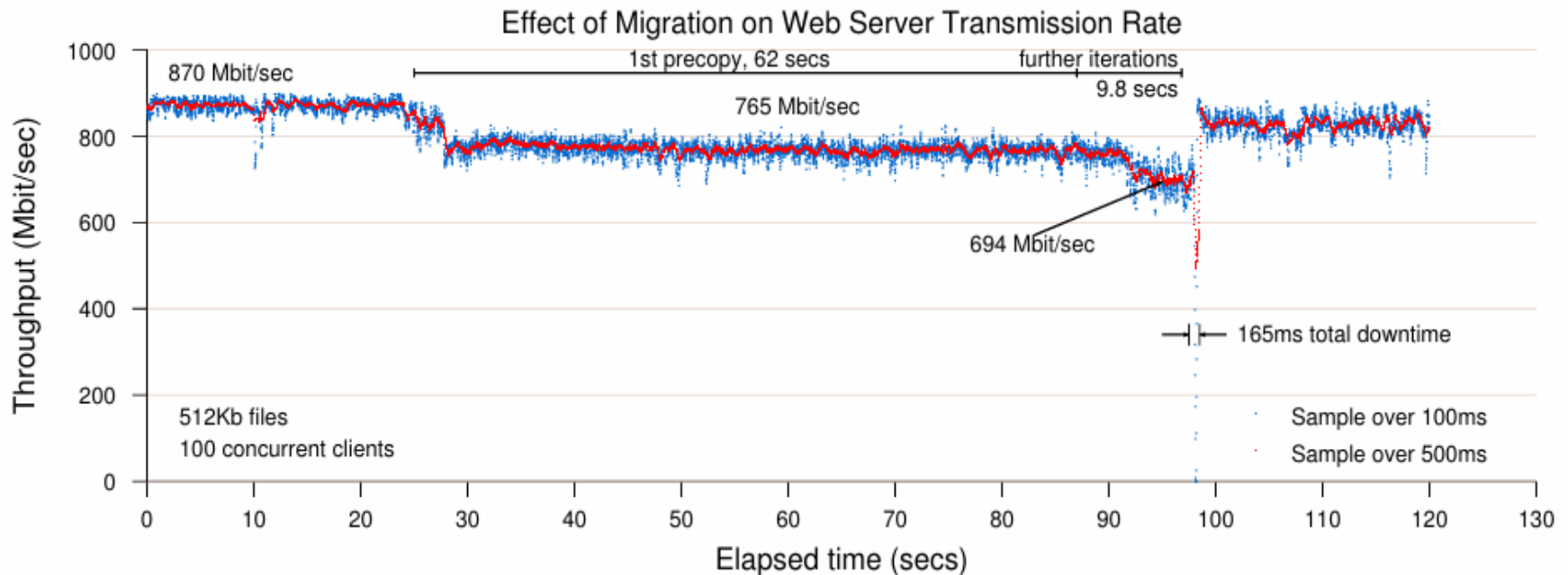
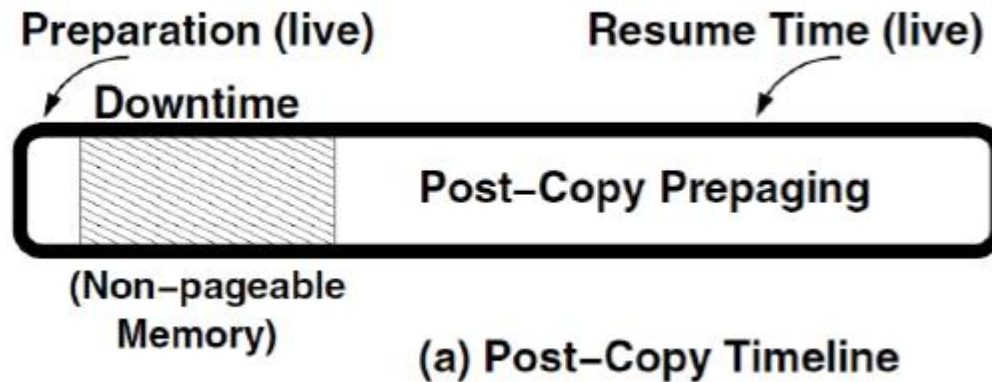


Figure 8: Results of migrating a running web server VM.

Post-copy Migration



1. Prepare target
2. Stop source VM, copy CPU context and minimum memory to target
3. Start VM at target, but without its memory!
4. Pull memory from source via demand paging
5. Memory access at target causes page fault, page fetched from source VM

Post-copy Migration



Advantage:

Each page transferred over the network only once.

Deterministic total migration time

Disadvantage:

Cold start penalty at the destination

If migration fails, then VM is lost.

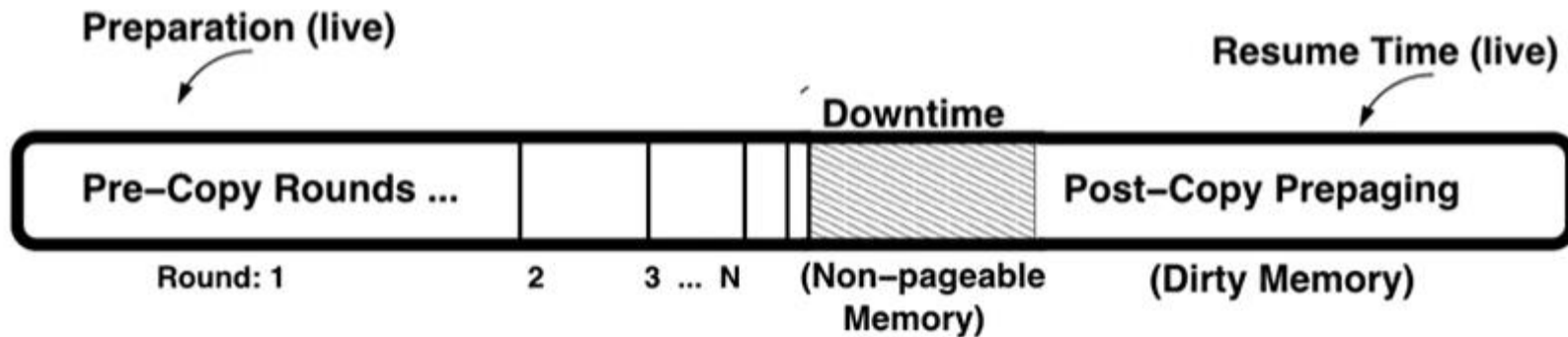
Optimizations on Post-copy Migration



- **Active pushing:** Source proactively pushes important pages, in addition to pulling pages via page faults
- **Pre-paging:** A “bubble” of pages around faulted page and proactively pushed are pulled, in anticipation of future accesses
- **Dynamic self-ballooning:** VM periodically frees up unnecessary memory and gives it back to hypervisor
 - Reduces memory footprint, speeds up page transfer
 - Performed carefully without hurting application performance
 - Can be used to optimize pre-copy migration as well
- **Hybrid:** one pre-copy round, followed by post copy

Slide text courtesy: IITB

Hybrid Pre/Post-copy



Combines the benefits & drawbacks of both

1. Perform one or more rounds of live pre-copy rounds
2. Pause VM and transfer execution state
3. Use post-copy to transfer any remaining dirty pages from source

Target machine Failure Handling



- **Pre-copy**
 - Simply aborts the migration, restart with another target
 - Latest state is on source only, so can recover
- **Post copy**
 - Source has stale memory, target has updated memory
 - If target crashes, cannot recover application data (unless some replication is performed)

Source Machine Failure Handling



- **Pre-copy**
 - Latest state was on source only
 - Simply aborts the migration, start fresh with the service.
 - If replication would have been there, service can be redirected to it and migration starts fresh from there.
- **Post Copy**
 - Source has stale memory, target has updated memory
 - Target may not have pulled all application data (unless some replication is performed)
 - Service has to be restarted

Pre-Copy Versus Post-Copy

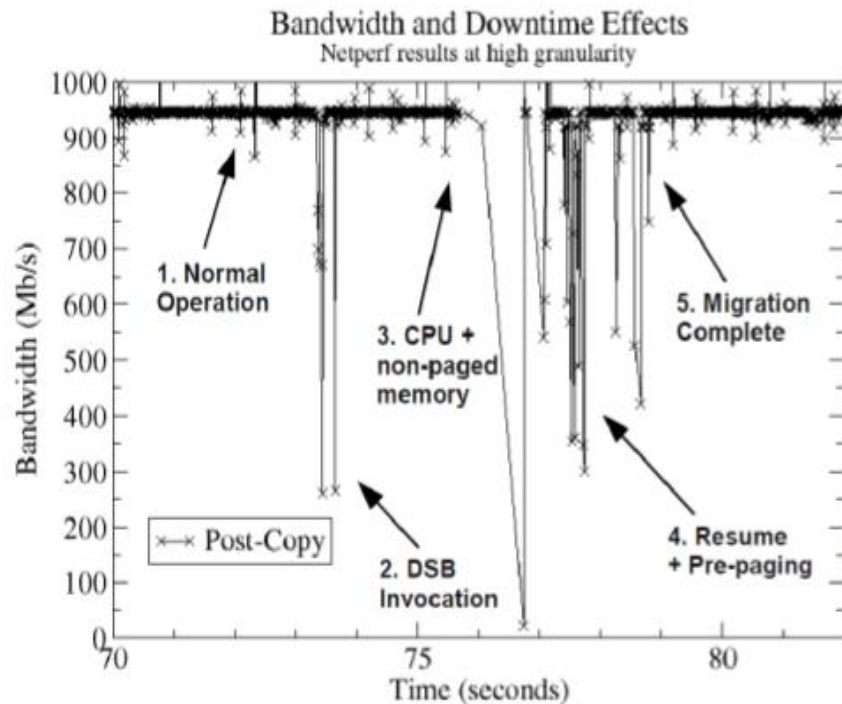


Figure 9. Impact of post-copy on NetPerf bandwidth.

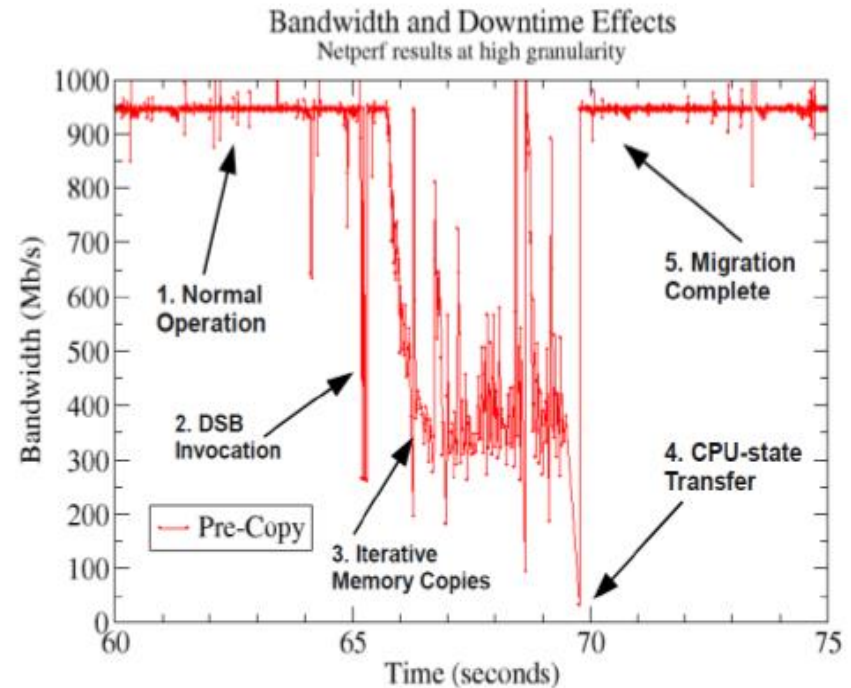


Figure 10. Impact of pre-copy on NetPerf bandwidth.

Image courtesy: "Post-Copy Based Live Virtual Machine Migration Using Adaptive Pre-Paging and Dynamic Self-Ballooning" Michael R. Hines and Kartik Gopalan

Pre-Copy Versus Post-Copy

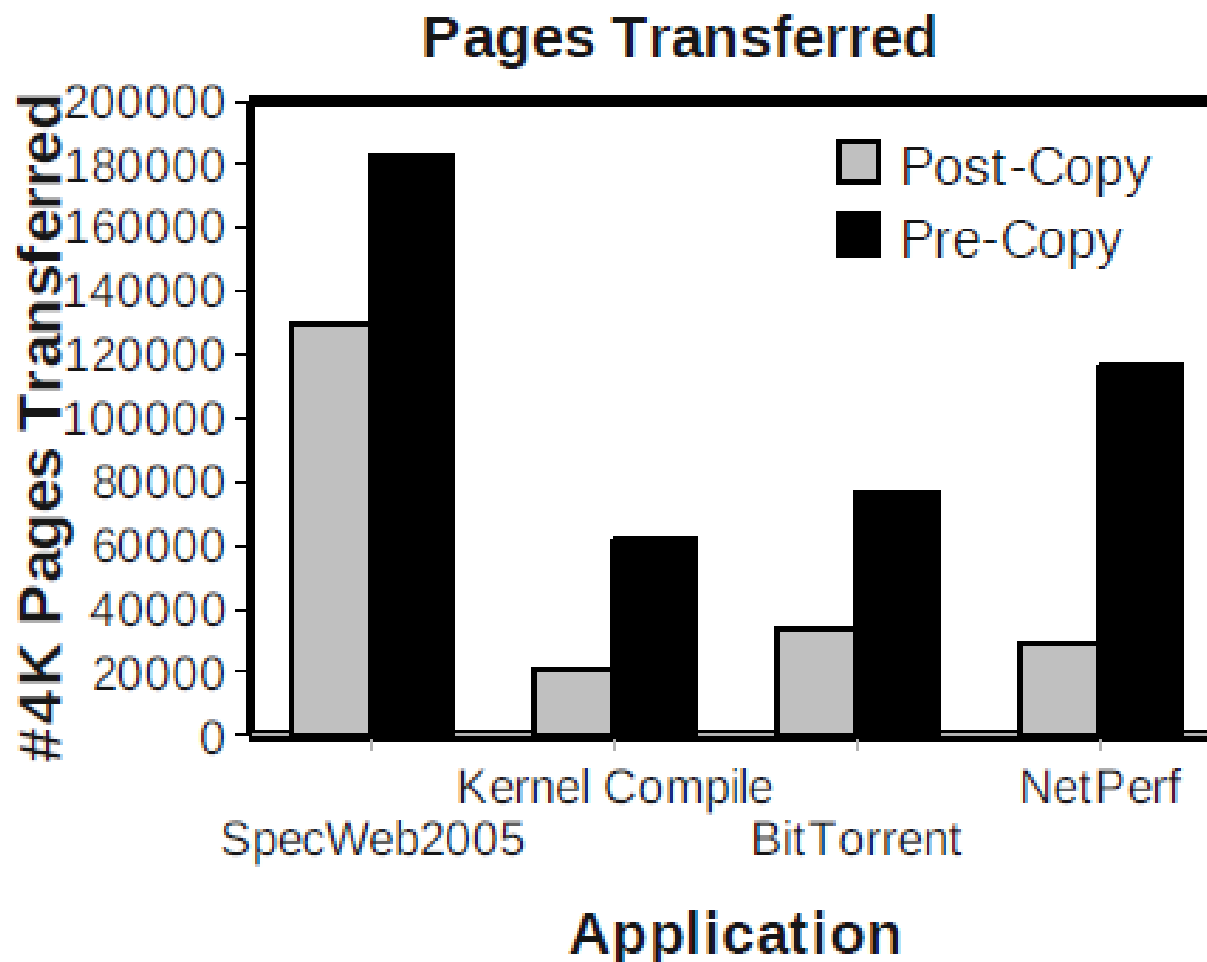
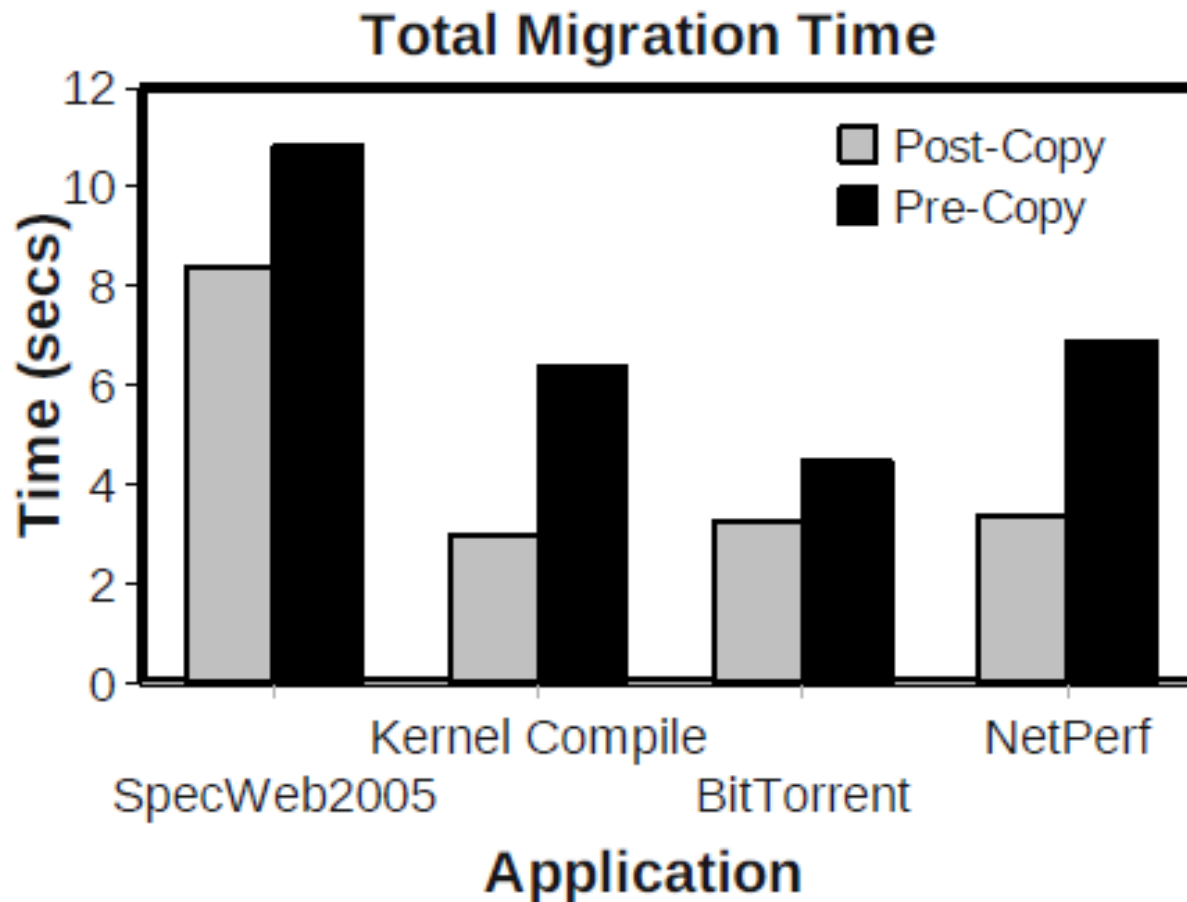


Image courtesy: "Post-Copy Based Live Virtual Machine Migration Using Adaptive Pre-Paging and Dynamic Self-Ballooning" Michael R. Hines and Kartik Gopalan

Pre-Copy Versus Post-Copy



Pre-Copy Versus Post-Copy

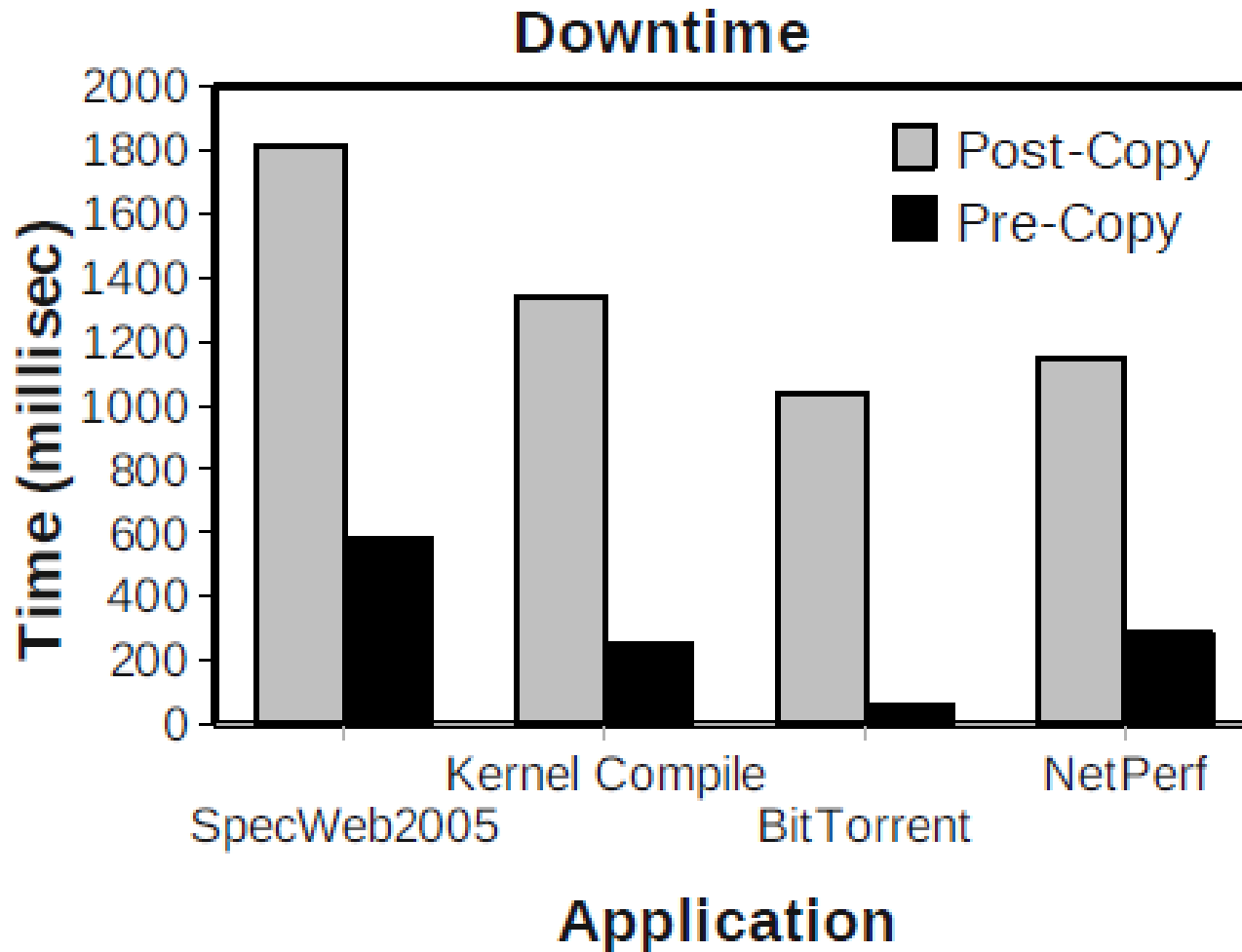


Image courtesy: "Post-Copy Based Live Virtual Machine Migration Using Adaptive Pre-Paging and Dynamic Self-Ballooning" Michael R. Hines and Kartik Gopalan

Pre-Copy Versus Post-Copy



- Post copy is better for memory-intensive applications
 - Longer downtime as compared to pre-copy,
 - But lower total migration time, fewer page transfers, lesser disruption to application
- Pre-copy suited for interactive application

Migrating Network Connections



Within a LAN

- The migrated VM carries its IP address, MAC address, and all protocol state, including any open sockets
- Backward (re)learning delay at the network switches
- Switches needs to re-learn the new location of migrated VM's MAC address
- **Solution:**
 - Send an unsolicited ARP reply from the target host.
 - Intermediate switches will re-learn automatically.
- Few in-flight packets might get lost.

Migrating Network Connections

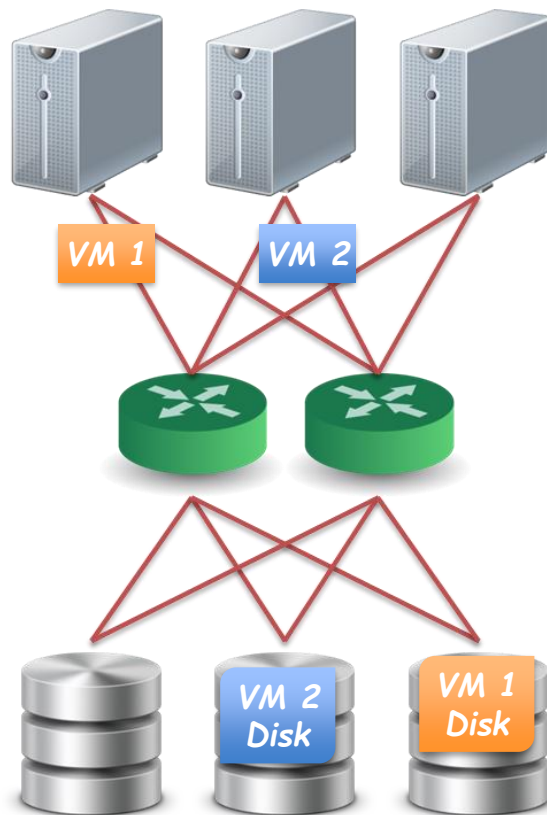


Across a WAN (Wide Area Network)

- Source and destination subnets may have different IP addresses.
- Active network connections may need to be tunneled via VPN or similar mechanisms.

Storage Migration

Many gigabytes of local disk image to migrate



Network Attached Storage

Storage Migration



For LAN

- Assume the storage is over the network and remains accessible from the new target machine.
- E.g. Network File System (NFS), or Network Block Device(NBD), or iSCSI etc.

For WAN

Disk image may need to be transferred.

- Can use pre-copy or post-copy for disk images
- Combined bandwidth saving optimizations such as compression, and/or de-duplication.

Summary

- VM Provisioning
- VM Migration
 - Cold/Regular
 - Hot/Live
- VM Live Migration
 - Pure Stop-and-Copy
 - Pre-Copy
 - Post-Copy
 - Hybrid of Pre-Copy and Post-Copy

References

- “Live Migration of Virtual Machines”, Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen, Eric Jul, Christian Limpach, Ian Pratt, Andrew Warfield, NSDI'05: Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation – Volume, 2005, doi: <https://dl.acm.org/doi/10.5555/1251203.1251223>
- “Post-Copy Based Live Virtual Machine Migration Using Adaptive Pre-Paging and Dynamic Self-Ballooning”, Michael R. Hines and Kartik Gopalan, VEE '09: Proceedings of the 2009 ACM SIGPLAN/SIGOPS international conference on Virtual execution environments, 2009, doi: <https://doi.org/10.1145/1508293.1508301>