Migration of Memory, Files, and Network Resources

Migration of Memory

- This is one of the most important aspects of VM migration
- Can be approached in any number of ways
- range of megabytes to a few gigabytes
- The ***Internet Suspend-Resume (ISR)*** technique exploits temporal locality
- **Temporal locality** refers to the fact that the memory states differ only by the amount of work done since a VM was last suspended before being initiated for migration.

Migration of Memory

- To exploit temporal locality, each file in the file system is represented as a tree of small sub files.
- A copy of this tree exists in both the suspended and resumed VM instances.
- the caching ensures the transmission of only those files which have been changed.
- The ISR technique deals with situations where the migration of live machines is not a necessity. Downtime is high.

File System Migration

- System must provide each VM with a consistent, location-independent view of the file system that is available on all hosts.
- To achieve this: “provide each VM with its own virtual disk which the file system is mapped to and transport the contents of this virtual disk along with the other states of the VM”.
- *not a viable solution*

File System Migration

- Another way is to have a global file system across all machines where a VM could be located
- all files are network accessible
- A distributed file system is used in ISR serving as a transport mechanism for propagating a suspended VM state

File System Migration

- The actual file systems are not mapped onto the distributed file system. Instead, the VMM only accesses its local file system.

File System Migration

- The relevant VM files are explicitly copied into the local file system for a resume operation and taken out of the local file system for a suspend operation
- However, this decoupling means that the VMM has to store the contents of each VM’s virtual disks in its local files, which have to be moved around with the other state information of that VM.

File System Migration

- In smart copying, the VMM exploits spatial locality.
- In these conditions, it is possible to transmit only the difference between the two file systems at suspending and resuming locations.

Network Migration

- A migrating VM should maintain all open network connections without relying on forwarding mechanisms on the original host or on support from mobility or redirection mechanisms
- VM must be assigned a virtual IP address known to other entities
- distinct from the IP address of the host machine where the VM is currently located

Network Migration

- Each VM can also have its own distinct virtual MAC address.
- The VMM maintains a mapping of the virtual IP and MAC addresses to their corresponding VMs.
- In general, a migrating VM includes all the protocol states and carries its IP address with it.

Network Migration

- If the source and destination machines of a VM migration are typically connected to a single switched LAN, an unsolicited ARP reply from the migrating host is provided advertising that the IP has moved to a new location.
- Although a few packets (transmitted) might be lost, no other problems with this method.

Network Migration

- Alternatively, on a switched network, the migrating OS can keep its original Ethernet MAC address and rely on the network switch to detect its move to a new port.

Live Migration

- Traditional migration suspends VMs before the transportation and then resumes them at the end of the process.
- By importing the pre-copy mechanism, a VM could be live migrated without stopping the VM and keep the applications running during the migration.

Live Migration

- Within a cluster environment where a network-accessible storage system, such as storage area network (SAN) or network attached storage (NAS), is employed only memory and CPU status needs to be transferred from the source node to the target node.

Live Migration

- Live migration techniques mainly use
  - the pre-copy approach: discussed earlier
    - In the pre-copy phase, although a VM service is still available, much performance degradation will occur

Live Migration

- Live migration techniques mainly use
  - A check pointing/recovery and trace/replay approach (CR/TR-Motion)
    - transfers the **execution trace file** in iterations rather than dirty pages, which is logged by a trace daemon.
    - Apparently, the total size of all log files is much less than that of dirty pages.
    - Total migration time and downtime of migration are drastically reduced

Network Migration

- Live migration techniques mainly use
  - strategy of post copy
  - all memory pages are transferred only once during the whole migration process and the baseline total migration time is reduced.
    - But the downtime is much higher than that of pre-copy due to the latency of fetching pages from the source node before the VM can be resumed on the target.
    - advent of multicore or many-core machines, abundant CPU resources are available

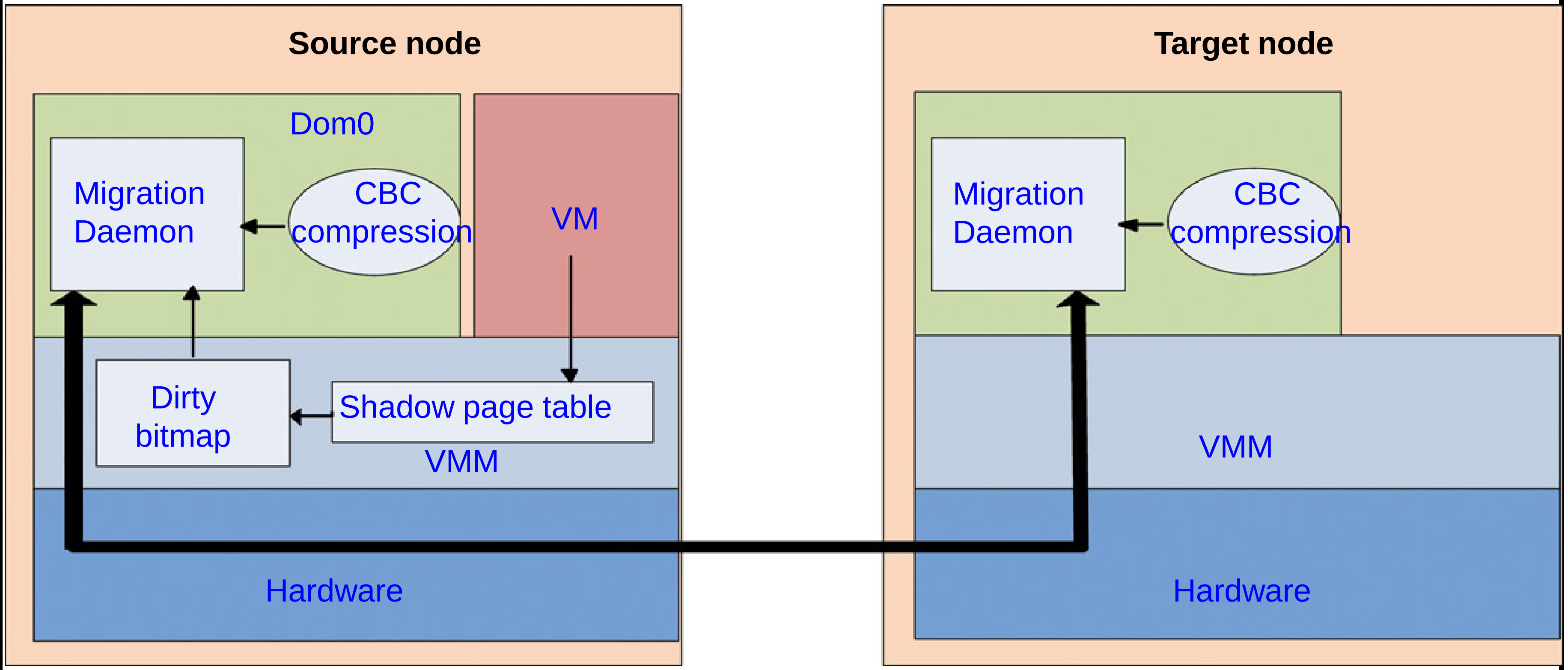
Live Migration of VM Using Xen

- Live migration transfers the working state and memory of a VM across a network when it is running.
- Xen also supports VM migration by using a mechanism called Remote Direct Memory Access (RDMA).
- RDMA speeds up VM migration by avoiding TCP/IP stack processing overhead.

Live Migration of VM Using Xen

- RDMA implements a different transfer protocol whose origin and destination VM buffers must be registered before any transfer operations occur, reducing it to a "one-sided" interface.
- Data communication over RDMA does not need to involve the CPU, caches, or context switches.
- This allows migration to be carried out with minimal impact on guest operating systems and hosted applications

Live Migration of VM Using Xen



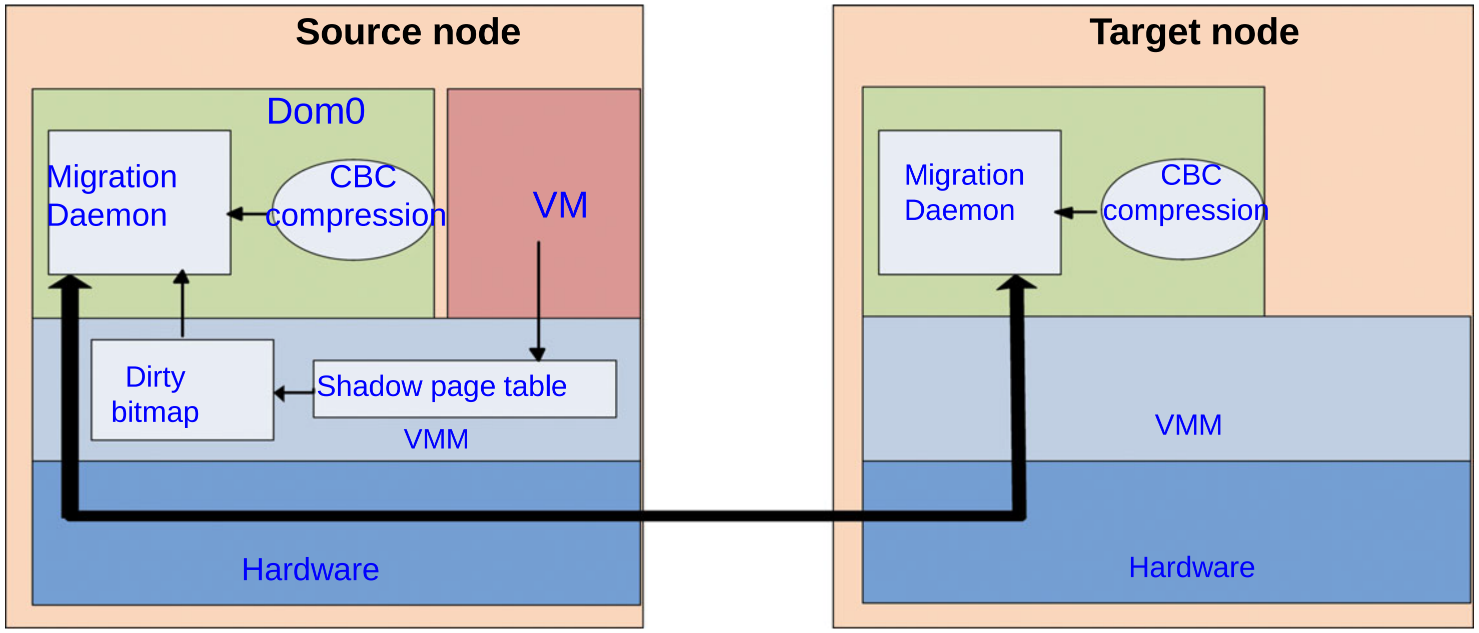
CBC: Characteristic Based Compression

Live Migration of VM Using Xen

- A single compression algorithm for all memory data is difficult to achieve.
- Therefore, there are compression algorithms to pages with different kinds of regularities.
- The structure of this live migration system is presented in **Dom0**.

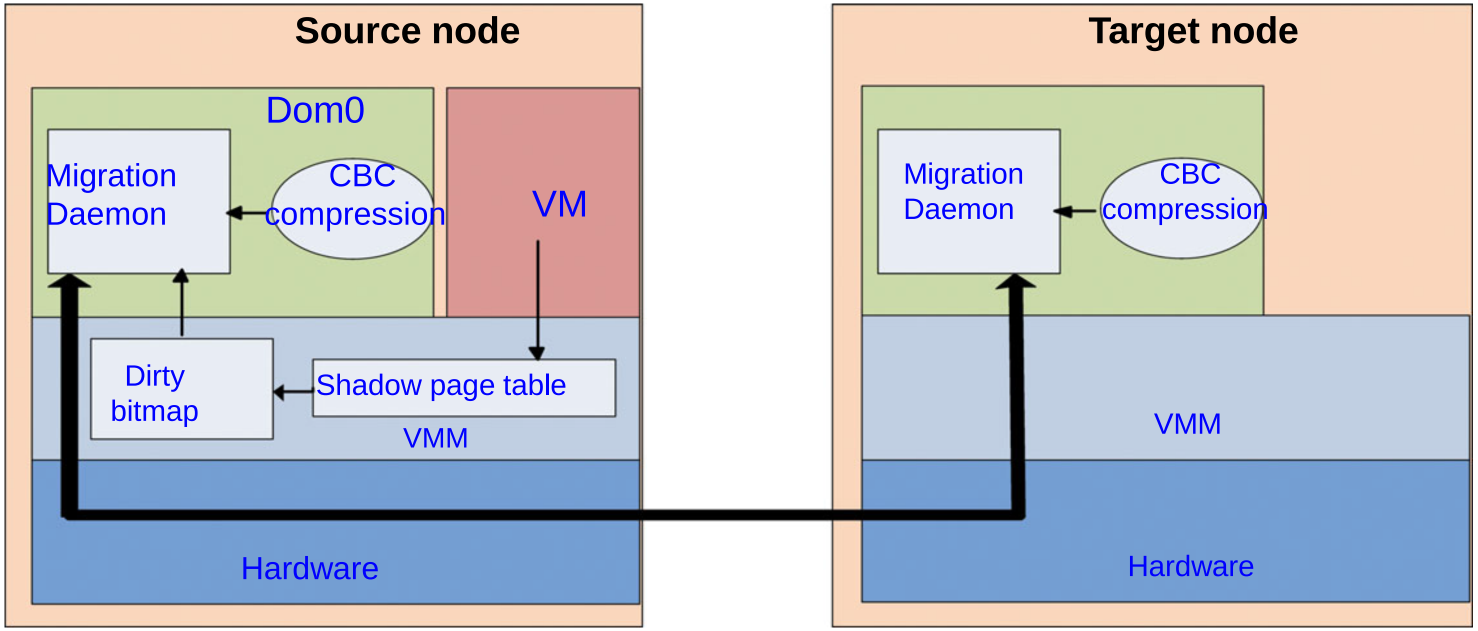
Live Migration of VM Using Xen

- Shadow page tables in the VMM layer trace modifications to the memory page in migrated VMs during the precopy phase. Corresponding flags are set in a dirty bitmap.



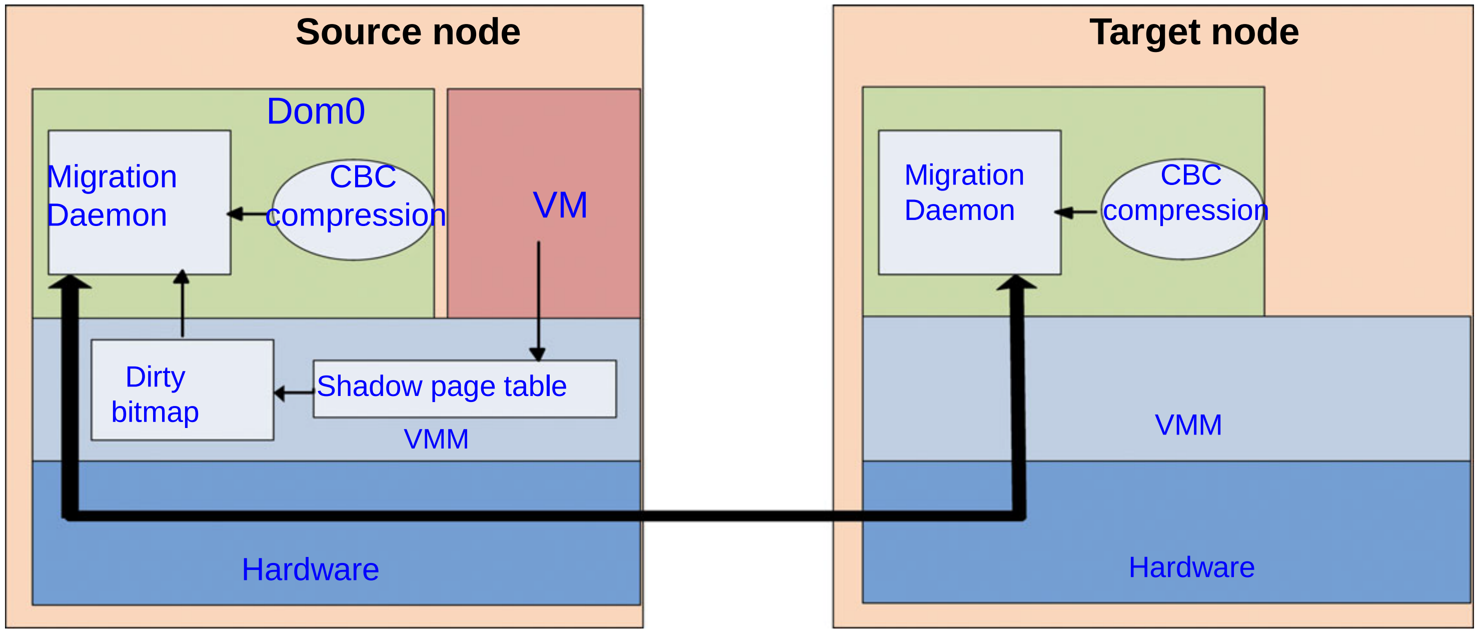
Live Migration of VM Using Xen

- At the start of each precopy round, the bitmap is sent to the migration daemon. Then, the bitmap is cleared and the shadow page tables are destroyed and re-created in the next round.



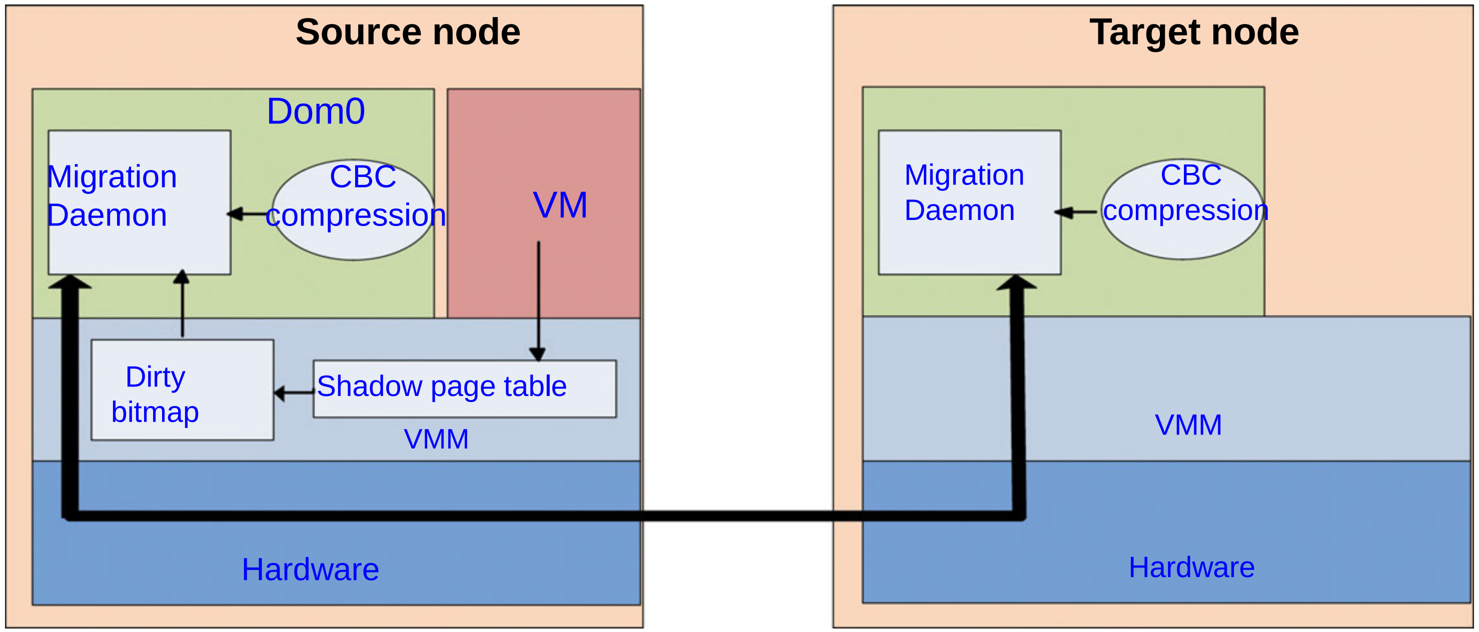
Live Migration of VM Using Xen

- The system resides in Xen's management VM. Memory pages denoted by bitmap are extracted and compressed before they are sent to the destination.



Live Migration of VM Using Xen

- The compressed data is then decompressed on the target.



Cloud OS for Virtualized Data Centers

Ref.: Distributed and Cloud Computing: From Parallel Processing to the Internet of Things Kai Hwang , Jack Dongarra , Geoffrey C. Fox.

Manager/ OS, Platforms,	License Resources Being Virtualized	Client API, Language	Hypervisors Used	Public Cloud Interface	Special Features
Nimbus, Linux, Apache v2	vSphere 4, Linux, Windows, proprietary	EC2 WS, WSRF, CLI	Xen, KVM	EC2	Virtual networks
Eucalyptus, Linux, BSD	Virtual networking	EC2 WS, CLI	Xen, KVM	EC2	Virtual networks
Open Nebula Linux, Apache v2	Management of VM, host, virtual network, and scheduling tools,	XML-RPC, CLI, Java	Xen, KVM	EC2, Elastic Host	Virtual Networks, dynamic provisioning
vSphere 4, Linux, Windows, Proprietary	Virtualizing OS for data centers	CLI, GUI, Portal, WS	VMware ESX, ESXi	VMware vCloud partners	Data protection, vStorage, VMFS, DRM, HA
OpenStack, Linus, Windows, OpenSource	Management of VM, host, virtual network, and scheduling tools, Virtual Cluster	Python API bindings and a CLI	KVM, LXC QEMU UML VMware Xen XCP Hyper-V	EC2 OpenStack cloud	Virtual Networks, dynamic provisioning VM migration Unlimited storage