Storage XenMotion: Live Storage Migration with Citrix XenServer®

Enabling cost effective storage migration and management strategies for enterprise and cloud datacenters

www.citrix.com

Table of Contents

1. Overview

Virtual machine (VM) agility is generally accepted as a crucial requirement for any virtualization solution. In 2007 the Burton Group, now part of Gartner, defined a set of criteria for enterprise grade virtualization, and included live migration of VMs as part of the core criteria. Since all versions of Citrix XenServer have had live VM migration, this was an easy component of the criteria for XenServer to satisfy.

What differentiates Live Storage Migration from Live VM Migration is that with Live Storage Migration the storage used for the virtual disks is moved from one storage location while the VM itself may not change virtualization hosts. In XenServer, Live VM Migration is branded XenMotion and logically Live Storage Migration became Storage XenMotion. Storage XenMotion is implemented using what is known as a shared nothing architecture. With a shared nothing architecture, the only assumption is the existence of a reliable high performance network between the source and destination, and this allows migration of VMs between XenServer resource pools.

2. Separating XenMotion from Storage XenMotion

For some readers, it may be helpful to review what differentiates a live migration event from a live storage migration event; including why neither of these solutions can be considered “zero-downtime”.

2.1 XenMotion Architecture

XenMotion implements a live VM migration strategy wherein the running state of a VM on one virtualization host is transferred to a second virtualization host without requiring the VM to reboot. In order for this technology to be effective, certain assumptions need to be made. The first crucial assumption is that the virtualization hosts share a common storage system, and that it is the running state which is being transferred, not the entire contents of the VM. Secondly, we need to ensure that the source and destination hosts implement a compatible CPU instruction set. Thirdly, we need to be aware of what the implications of a VM moving mean to the network.

2.1.1 Compatible CPU Instruction Set

In the early days of Intel VT based virtualization, resource pools were typically constructed using CPUs from identical processor families. This was done in large part to ensure that the running machine state for a VM on one host could actually run on a second host within the same pool. Failure to ensure compatible hardware could easily result in VMs crashing or worse silent corruption of the VM state when CPU instructions differed slightly between processor families or even CPU steppings. Over time technologies like Intel FlexMigrate eased restrictions on processor family similarities by allowing the processors in a pool to create an instruction equivalency mask. The equivalency mask effectively creates a common set of CPU instructions which all hosts in a resource pool can execute thus avoiding the potential for VM crashes.

2.1.2 Transferring the VM Running State

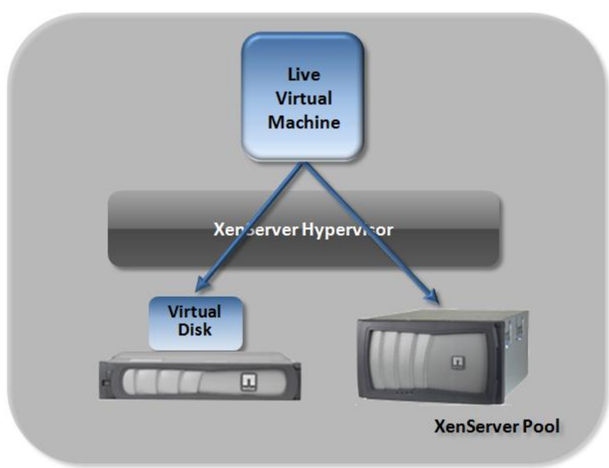
When a XenServer administrator, or workload management software, decides that a running VM should be migrated within its resource pool, XenMotion starts. XenMotion is a background process which allows a VM to continue accepting and processing user and network requests while the VM is transferred to another host. This process is not instantaneous, and the length of time it takes to complete is based on a number of factors including available network bandwidth, guest VM activity and host activity on both the source and destination. XenMotion does require either a paravirtualized (PV) VM, or a VM running the XenServer PV tools. The requirement for either a PV operating system or PV tools is dictated by a requirement to migrate the entire running state, including the state of the various virtual devices attached to the system without requiring a reboot of the VM.

2.1.3 Network Implications

For those people unfamiliar with networking, live VM migration has come to be defined as including a zero-downtime attribute. The desire is always to have the migration complete without any requirement for users to reconnect in any way to the running VM. In reality all live VM migration events will experience a minor outage, and this outage is due to a need for the underlying network to understand where the VM's MAC and IP addresses have moved to in the network.

2.2 Storage XenMotion Architecture

Unlike XenMotion which operates on the VM running state, Storage XenMotion operates solely on the virtual disk implementation (vdi) which supports each of the disks available to the VM. XenServer supports the use of host local storage, direct attached storage, NFS and iSCSI or HBA connected block based storage solutions for storage repositories (SR). While the actual storage implement differs between storage types, the core file format remains the same. XenServer uses the Microsoft VHD specification for its vdi implementations, and included in that specification are options for snapshot chains.



When an administrator requests that Storage XenMotion should move a given vdi from one SR to another, XenServer performs a snapshot on the vdi. VM operations are seamlessly directed to this snapshot. Concurrently the snapshot is written to the new SR resulting in a mirrored write which ensures that write operations are synchronized between the source and destination SRs. Once mirroring is established, the root node of the vdi begins to transfer. Since all VM activities are occurring on the mirrored snapshot, storage migration will complete once the root node has been completely migrated. One of the core requirements for Storage XenMotion was to be tolerant of storage configurations. This includes scenarios where the virtualization administrator is unaware of the storage provisioning model and as a result could accidentally migrate a vdi for which there is insufficient storage on the destination. With Storage XenMotion, if any error occurs in the migration event, the migration event is simply aborted and the destination objects cleaned up. No interruption to VM operations occur in such situations. The result of this model is a predictable and reliable storage migration solution with minimal performance impact.

2.2.1 Network Impact

In a XenServer environment, networks can be either management networks or VM guest networks. Management networks are used for all management operations, inter host communications, and for storage traffic. By default all XenServer installations have a single primary management network, though additional secondary management networks are typically created to isolate storage traffic from VM traffic.

Since storage migration requires the entire contents of the virtual disk to be transferred from one storage solution to another, the impact of this transfer needs to be considered. The most important item to be aware of is that both XenMotion and Storage XenMotion occur over XenServer management interfaces. In versions prior to XenServer 6.1, to limit the impact of migration on VM IO, this transfer was always over the primary management interface. Starting with XenServer 6.1, administrators initiating XenMotion and Storage XenMotion operations can specify which management interface transfers should occur over. Through the use of multiple management interfaces, the virtual disk transfer can occur with minimal impact on both core XenServer operations and VM network utilization.

2.2.2 Storage Migration Security

In traditional virtualization environments with dedicated shared storage networks, sensitive disk data is transferred across known networks with clearly defined restrictions built into them. Storage Migration changes this model to potentially allow disk data to be transferred over general purpose networks, including networks which may have configurations that mirror traffic. With data retention policies and compliance considerations being different on general purpose networks when compared to dedicated storage networks, Storage XenMotion was designed to fully encrypt the VM migration traffic. This encryption practice was also applied to traditional XenMotion to ensure that running VM state was also protected from network monitoring practices.

3. Shared Nothing Migration – Meeting Cloud Agility Needs

While live storage migration has been available from multiple virtualization vendors for some time, traditional storage migration has been limited to a single resource pool and required shared storage. This limitation is rarely relevant in small business or enterprise environments, but when cloud agility is required, the nature of resource pools and cloud architectures makes traditional live storage migration less viable.

3.1 Cloud Architectures and the Economics of Service Delivery

While traditional virtualized enterprise datacenters have been built around technologies like resource pools and high performance shared storage with multiple data paths having implemented features like live VM migration and a variety of virtualization high availability strategies, cloud deployments are very different. For most cloud operators, the costs of shared storage have prevented wide adoption of the traditional agility options found in the enterprise. This has led many cloud operators to adopt a model using only local storage for running VMs, and this model is fully supported by leading cloud orchestration solutions such as Citrix CloudPlatform, Apache CloudStack and OpenStack. When leveraging a local storage only option, you naturally give up the potential of live VM migration and rely on the agility and resiliency of the application being hosted to ensure adequate uptime. When shared storage is used by cloud operators, it tends to be implemented using low cost options such as NFS and resource pool sizes also tend to be small. Cloud operators implementing a shared storage model are doing so with an eye on providing the benefits of live VM migration and virtualization host failure protection as value add features to their customer base. Despite using shared storage, the reality of cloud scale deployments dictate that true VM agility requires a live migration option which doesn't rely on shared storage or resource pools and effectively implements a shared nothing storage migration model.

4. Conclusion

Virtual machine agility is one of the key features in a successful virtualization strategy. While traditional live migration options offer one option, they only address part of the problem. A reality of datacenter operations is that storage does fill up, and that storage system upgrades are required. Storage XenMotion, available in XenServer 6.1 Advanced Edition, effectively addresses these common datacenter use cases while supporting a number of additional scenarios as outlined in the following table.

Scenario	XenMotion	Storage XenMotion	Shared Nothing XenMotion
Live migration of a VM within a resource pool	Yes	Not Required	Not Required
Live migration of a VM disk across shared storage within a resource pool	Not Required	Yes	Not Required
Live migration of a VM disk from one storage type to another within a resource pool	Not Applicable	Yes	Not Required
Live migration of a VM disk to or from local storage on a XenServer host within a resource pool	Not Required	Not Required	Yes
Live migration of a VM from one resource pool to another resource pool	Not Required	Not Required	Yes
Live migration of a VM from one CPU vendor to a different CPU vendor	Not Available	Not Available	Not Available

About Citrix

Citrix Systems, Inc. (NASDAQ:CTXS) is the company transforming how people, businesses and IT work and collaborate in the cloud era. With market-leading cloud, collaboration, networking and virtualization technologies, Citrix powers mobile workstyles and cloud services, making complex enterprise IT simpler and more accessible for 260,000 enterprises. Citrix touches 75 percent of Internet users each day and partners with more than 10,000 companies in 100 countries. Annual revenue in 2011 was $2.21 billion. Learn more at www.citrix.com.