香 港 大 學
THE UNIVERSITY OF HONG KONG

A survey on VM migration over WAN

Linquan Zhang

Department of Computer Science
The University of Hong Kong

May. 2011

Outline

香 港 大 學
THE UNIVERSITY OF HONG KONG

Cloud

- Cloud is a promising computing paradigm, which allows hosting of multiple services on a globally shared resource pool where resources are allocated to services on demand.
- Typical cloud platforms: Amazon EC2, Google App Engine
- Typical applications: Dropbox, foursquare, notepad.cc

香 港 大 學
THE UNIVERSITY OF HONG KONG

Virtualization

- In cloud computing, we can consider it as the creation of a virtual version of hardware platform
- To provide isolated domains for guest programs, as if they were running on a separate system
- To support different applications running on a cloud
- To improve flexibility of resource provisioning
- Major virtualization platforms: Xen, KVM, Microsoft Hyper-V Server, VMware VirtualCenter

What is virtual machine(VM) migration ?

- Virtual machine migration takes a running virtual machine and moves it from one physical machine to another. Usually, migration is transparent to the guest operating system and applications
- VM migration is supported by major virtualization platforms, such as Xen, Microsoft Hyper-V Server, VMware VMotion
- In this presentation, we only consider live VM migration

The motivation of VM Migration

- To improve global system utilization by load balancing across physical machines
- To improve system serviceability and availability by moving applications off machines that need servicing or upgrades.
- To avoid many difficulties faced by process-level migration approaches. E.g. the problem of residual dependencies.
- To provide continuous services while migrating (Live migration). E.g. we can migrate a streaming media server without requiring clients to reconnect

Details about VM Migration

There are three kinds of states that need to be dealt with when migrating a VM:

- The virtual device state including the state of the CPU, the motherboard, networking and storage adapters, and graphics adapters
- External connections with devices including networking, USB devices, SCSI storage devices, and removable media such as CD-ROMs
- The VM's physical memory

VM Migration for Cloud

When it's specified to the case of cloud computing:

- The state of the CPU, networking, and storage devices
- External connections with devices, particularly networking connections
- The VM's physical memory

Requirements of VM migration

There are two metrics we need to minimize:

- Downtime: the period during which the service is unavailable due to there being no currently executing instance of the VM
- Total migration time: the duration between when migration is initiated and when the original VM may be finally discarded.

However, there are trade-offs between these two requirements.

VM migration over LAN

There are some assumptions:

- Disk image is stored on a network-attached storage(NAS), so that migration only moves memory state as well as network state
- The physical machines are in a cluster. They are connected via LAN, which has high speed and low latency.

VM migration over LAN

- Consider a virtual machine to encapsulate access to a set of physical resources, we focus on the physical resources used, particularly on memory, network and disk
- Due to previous assumptions, there is no need to pay attention to migrating disk state. NAS is uniformly accessible from all host machines in the cluster
- We need to address two kinds of migration: memory and network migration.

Memory migration

There are three main approaches to address this issue:

- Pure stop-and-copy: halt the original VM, copy all pages to the destination, and then start the new VM.
- Pure demand-migration: a short stop-and-copy phase transfers essential kernel data structures to the destination, then destination VM is started, and other pages are transferred across the network on first use.
- ★ Pre-copy: memory pages are iteratively copied from the source machine to the destination one, all without stopping the execution of the migrated VM.[1]

Network migration

For network resources, we want a migrated OS to maintain all open network connections. A migrated VM will include all protocol state (e.g. TCP), and will carry its IP address with it.

- In a cluster environment, the network interfaces of the source and destination machines typically exist on a single switched LAN
- Generating an ARP reply from the migrated host to advertise that the IP has moved to a new location.
- Only a very small number of in-flight packets may be lost.

Migration timeline

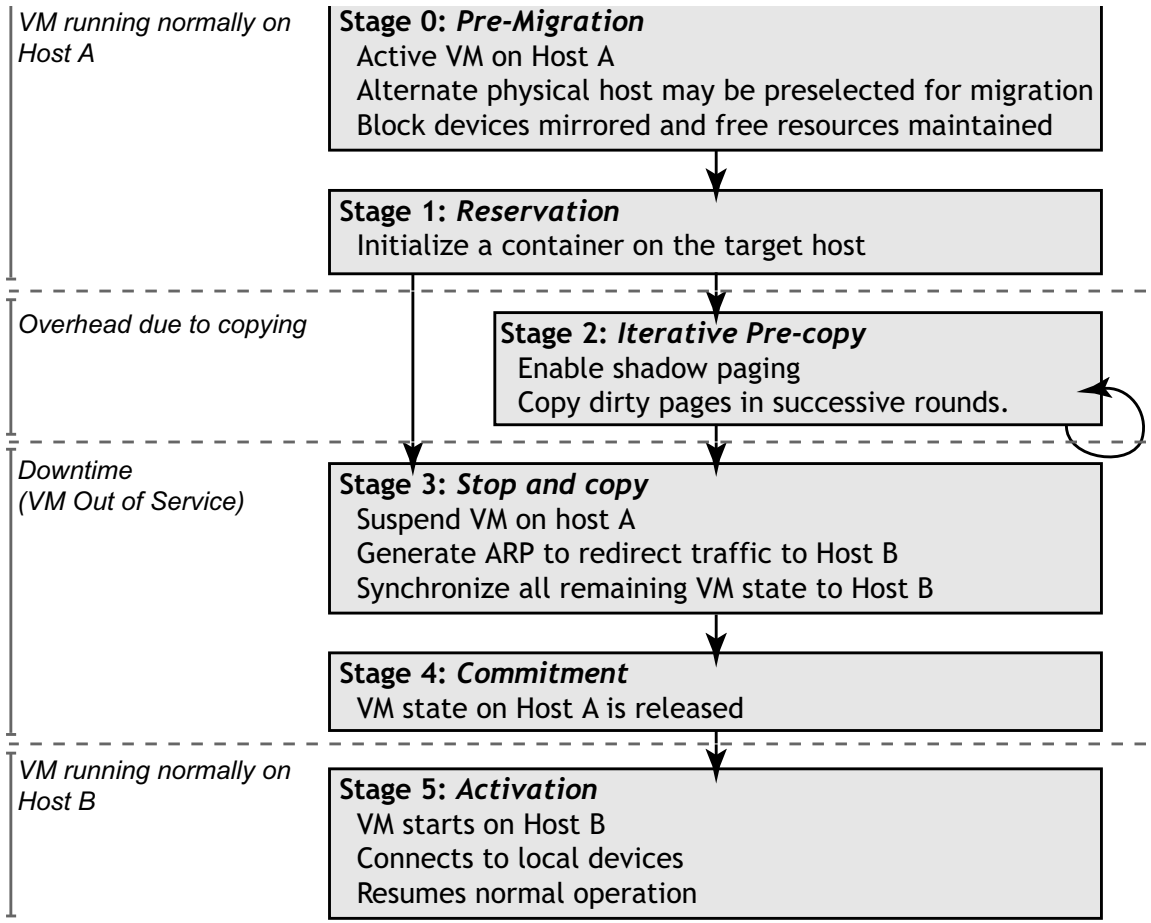

Figure: Migration timeline[1]

Other proposed solutions

- Huang et al. have proposed to optimize migration within the LAN by exploiting fast interconnects that support remote memory access technology[2]
- Jin et al. have proposed to use memory compression algorithms to optimize migrations[3]
- Breitgand et al. have developed a model based approach to determine when to stop iterating during a memory migration[4]

VM migration over WAN

Two key differences:

- Lower bandwidth and higher latency in WAN link.
- The existing LAN-based VM migration algorithms assume that disk state is stored on a NAS. However, for WAN-based VM migration, this assumption does not hold since the NAS may not span multiple data center sites

We need to improve existing LAN-based VM migration algorithms.

Existing issues

- Migrating network transparently (IP address keeps unchanged, TCP connections move over) over a WAN is a major challenge
- Low bandwidth leads to poor performance and poor user experience
- The disk state is required to move to destination physical machine. However, synchronizing disk state is high costly, and consumes a lot of bandwidth

Existing solutions

The existing research investigating migration of VMs over WAN has focused on either storage or network concerns.

- Bradford et al. describe a WAN migration system focusing on efficiently synchronizing disk state during the migration [5]
- Riteau et al. use content based addressing to detect redundancy across multiple hosts at the destination site during VM migration [6]

Existing solutions

- The VM Turntable Demonstrator shows a VM migration over intercontinental distances with latencies of nearly 200 msec; they utilize gigabit lightpath links to find that the increasing latency has less impact on performance than bandwidth [7]
- Harney et al. propose the use of Mobile IPv6 to reroute packets to the VM after it is moved to a new destination [8]
- ★ Wood et al. present a set of optimizations that minimize the cost of transferring storage and virtual machine memory during migration. Meanwhile, they propose Virtual Cloud Pool(VCP) to support network reconfiguration [9]

A sample solution

A set of optimizations

- Smart Stop and Copy: to reduce the number of unnecessary iterations and to pick a stopping point that minimize pause time
- Content Based Redundancy: to save bandwidth, and to eliminate the redundant data while transferring VM memory and disk image
- Using Page Deltas: keep a cache of previously transmitted pages, and then only send the difference between the cached and current page if it is retransmitted

A sample solution

Virtual Cloud Pool

- It is an abstraction that allows CloudNet to seamlessly connect geographically separate servers, it can be considered as a form of network virtualization
- To provide an illusion of a single logical pool of resources connected over a LAN

Conclusion

There still exists many challenges on VM migration over WAN.

- How to transfer memory state efficiently, as well as disk state
- How to guarantee the security while migrating over WAN
- How to reconfigure network quickly after migrating
- How to minimize the performance degrade while migrating
- ...

Reference

[1] C. Clark, K. Fraser, S. Hand, J. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield. Live migration of virtual machines. In Proceedings of NSDI, May 2005.
[2] W. Huang, Q. Gao, J. Liu, and D. K. Panda. High performance virtual machine migration with RDMA over modern interconnects. In Proceedings of the 2007 IEEE International Conference on Cluster Computing, pages 11 – 20. IEEE Computer Society, 2007.
[3] H. Jin, L. Deng, S. Wu, X. Shi, and X. Pan. Live virtual machine migration with adaptive memory compression. In Cluster, 2009.
[4] D. Breitgand, G. Kutiel, and D. Raz. Cost-aware live migration of services in the cloud. In Proceedings of the 3rd Annual Haifa Experimental Systems Conference, SYSTOR ' 10, New York, NY, USA, 2010. ACM.
[5] R. Bradford, E. Kotsovinos, A. Feldmann, and H. Schi ¨ oberg. Live wide-area migration of virtual machines including local persistent state. In Proceedings of the 3rd international conference on Virtual execution environments, pages 169 – 179, San Diego, California, USA, 2007. ACM.
[6] P. Riteau, C. Morin, and T. Priol. Shrinker: Efficient Wide-Area Live Virtual Machine Migration using Distributed Content-Based Addressing. Research Report RR-7198, INRIA, 02 2010.
[7] F. Travostino, P. Daspit, L. Gommans, C. Jog, C. de Laat, J.Mambretti, I. Monga, B. van Oudenaarde, S. Raghunath, and P. Y. Wang. Seamless live migration of virtual machines over the MAN/WAN. Future Generation Computer Systems, Oct. 2006.
[8] E. Harney, S. Goasguen, J. Martin, M. Murphy, and M. Westall. The efficacy of live virtual machine migrations over the internet. In Proceedings of the 3rd VTDC, 2007.
[9] T. Wood, P. Shenoy, K. K. Ramakrishnan, and J. V. der Merwe. CloudNet: Dynamic Pooling of Cloud Resources by Live WAN Migration of Virtual Machines. In Proceedings of VEE'11, March. 2011

Q&A

Thanks for your attention!