Virtualization
and
Disk Performance

Introduction:

Given that virtualization technologies have many specific applications this paper will begin by first presenting definitions.

Definition: Virtualization

Essentially to virtualize something means to make something that doesn't actually (physically) exist *appear* to exist. Think of the context in *virtual reality*. Let's make a quick example of something everyone in IT is familiar with, a PC with 4 logical volumes (C, D, E, and F). In reality that desktop has one physical disk drive partitioned into 4 volumes. A logical volume is in this case a virtual drive.

Next we'll define two popular modern applications of virtualization technology.

Definition: Server Virtualization[1] / Virtual Machine

Server virtualization describes the creation of one or more virtual instances of a "guest" operating system either on top of a "host" operating system (Hosted Architecture) or directly on top of a specialized software layer called a hypervisor (Hypervisor Architecture).

Image1 (virtual machines)

In either architecture, the host system's virtualization of other operating systems is accomplished by software, proprietary to the vendor (e.g. Virtual Server™, VMware™, Virtuozzo™), which resides between the physical hardware (CPU, memory, etc) and the "guest" operating systems. Each guest or host operating system runs its own applications independently, as if it were the only system operating on the hardware.

In a host/guest environment, each instance of a guest operating system stores a file called a virtual disk (e.g. .vhd, .vmdk) on the host system. This is a very common implementation of machine virtualization today.

Hypervisor architecture removes the requirement for a "host" system. With a hypervisor, virtual machines run on a thin layer of hardware abstraction software. That software layer, the hypervisor, addresses hardware communications for all the virtual systems on that machine. Hypervisor represents the future of virtual machine technology.

---

[1] For the general purposes of this document, server virtualization also includes PC (workstation) virtualization.

Definition: Storage Virtualization
Storage virtualization involves the creation of a usually very large, logical-pool of data that, via software, appears to be physically located all on one server. In actuality, that data may be located across hundreds of physical disks spread across dozens of servers. This is the concept implemented by Storage Area Networks (SAN). For peak performance these storage pools require automatic disk defragmentation just the same as a single hard drive would. Automatic defragmentation is implemented from server(s) that manage the respective logical disk volumes.

Image2 (storage virtualization)

Our last definition is a broad explanation of disk fragmentation.

Definition: Disk Fragmentation
Disk fragmentation, is the condition in which pieces of individual files and free space on a disk are not contiguous, but rather broken up and scattered around the disk. This requires the hard drive to locate all the fragments of a file. The collection of file fragments from numerous places instead of just one causes file access to take significantly longer than it should. File writes into fragmented free space, also take longer and can increase the likelihood of newly created files fragmenting.

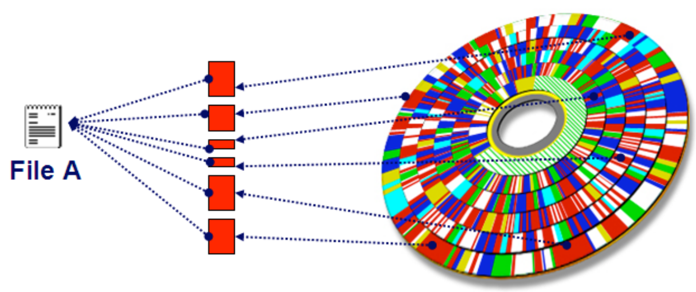

Image3 (disk fragmentation)

The affect of disk fragmentation is slower system performance, increased I/O overhead, and more severe cases, compromised reliability resulting in phenomena such as application and system hangs and crashes.

Overview:

Depending on your perspective, *virtualization's* purpose is to afford divergence and convergence. It affords the division of logical objects that should be separated, and/or the consolidation of objects that should be grouped together.

The technology's recent explosion coincides with the trend of consolidating systems on to fewer, but more powerful hardware. With more robust hardware, consolidation makes cost-effective sense. And given consolidation for the purpose of reduced management overhead and more efficient hardware utilization, virtualization makes a great deal of sense.

The purpose of *defragmentation* is to consolidate file fragments into a single extent, increasing access speed, and to reduce free space fragments to a small handful of larger chunks.

Virtualization does have its dangers, as it incurs greater stress on physical resources. While under utilization of CPU may be a driving factor to virtualize servers, other hardware resources may become overtaxed. Given that a host system has limited ability (depends on application) to page memory used by the guest systems, the most recognized bottleneck to address is physical memory (RAM). Options to programmatically alleviate memory bottlenecks incur performance issues when the disk is re-introduced. Another major component and perhaps less acknowledged is the disk subsystem. In many cases, depending the purpose and application of the guest/virtual systems, the disk bottleneck will be the most significant barrier to performance.

The remainder of this paper will discuss the increased importance of disk performance.

The Disk is the Weak Link:

CPUs and memory operate orders of magnitude faster than mechanical hard drives. The slower the disk, the slower the entire system will be.

While these facts are well known to industry professionals, it deserves re-iteration as the issue becomes manifest when those *disks* are asked to do more. Such is the case with virtualization, where the given hardware has to support numerous simultaneous operating systems.

Another vital factor to consider is that server virtualization can compound disk fragmentation; and as we covered earlier, disk fragmentation slows disk performance.

Typically fragmentation occurs on logical disk drives, and by device drivers is translated to physical sectors on a disk. It can be demonstrated as pieces of a file located in a non-contiguous manner (Image4). In the case of virtual systems, the logical volume is masked by the technology; known as a virtual disk. These virtual disks reside on logical disks in the form of container files. Those virtual disk files can fragment just as any other file can resulting in what amounts to a “logically” fragmented virtual hard disk (Image5), which still has typical file fragmentation contained within it.

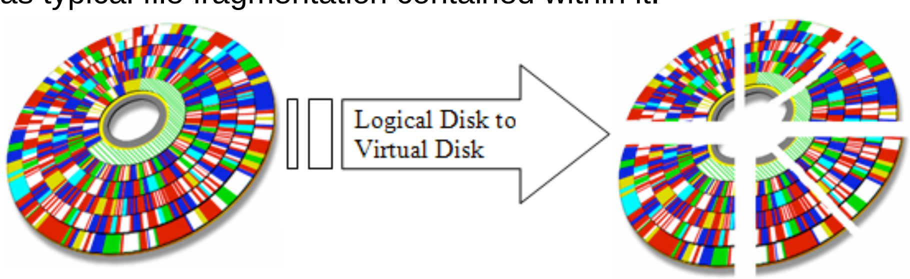

Image 4 (fragmented drive) ***Image 5 (fragmented virtual disk)***

The picture represented in Image 5 would appear in a defragmentation analysis report's "Fragmented Files" list run from a host Windows operating system as "VirtualServer1.vhd, 4GB in size, in 6 pieces".

This equates to hierarchical fragmentation or more simply fragmentation-within-fragmentation. The black lines in Image 6 (below left) represent disk I/O mappings of the virtual disk file fragments to the host system in a Hosted Architecture. The smallest unit of data access in a virtual machine is typically 128 sectors, or 64KB. Therefore if these access units (called grains in VMware) are fragmented, performance suffers.

Image 7 (below right) depicts a fragmented file ("Fragmented Word doc in Virtual Server.doc") residing on a virtual disk, which in turn exists as a fragmented file on the host operating system. The current design of software-based server virtualization requires the host system capture and process any disk I/O generated by guest operating systems, adding an additional layer in the I/O processing stack.

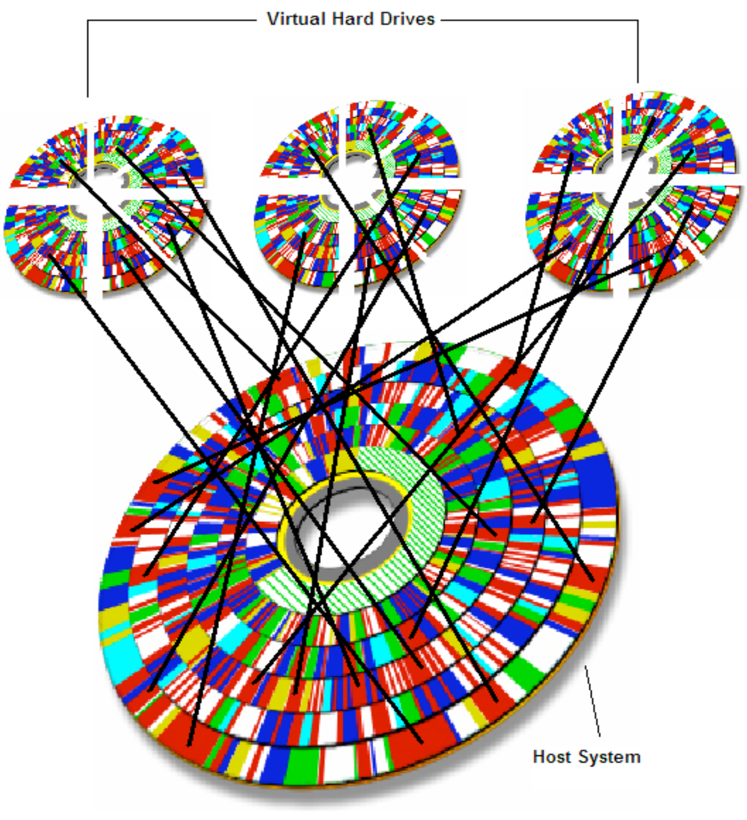

Image 6 (Hierarchical disk fragmentation)

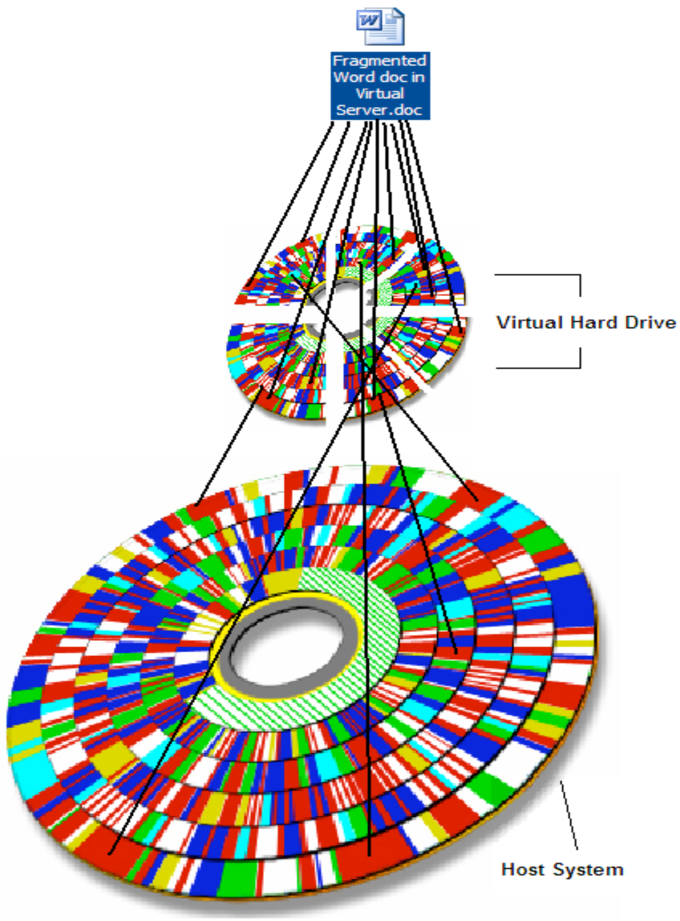

Image7 (tiered disk I/Os to a fragmented file in a fragmented virtual disk

Machine Virtualization Architectures and I/O:

Given either of two predominant virtualization architectures (Hosted or Hypervisor) remember that the virtual machines are emulating hardware and may not emulate the exact specifications. For example, a high-end video card may not be emulated in a host system with all the advanced capabilities.

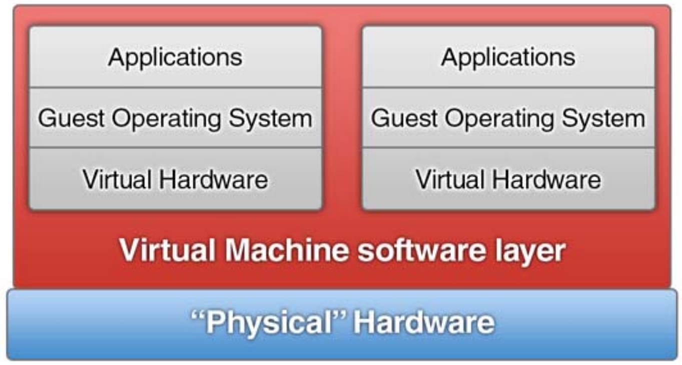

Image 8 (Hypervisor Virtualization)

The Hypervisor architecture (right) removes the requirement for a host operating system and improves overall virtual systems performance.

As demonstrated earlier, Disk I/O's generated from virtual systems (Hosted Architecture) can suffer from increased software stack processing. This means that disk I/O has to go up and down software layers that abstract the physical hardware. In a Hosted Architecture, a low level disk request in a guest system is translated into a user–level call in the host system. With the likely loss of disk caching at the guest level (hardware support consideration), and limited queuing ability, this process will not be as speedy as a direct physical hardware call by the host system.

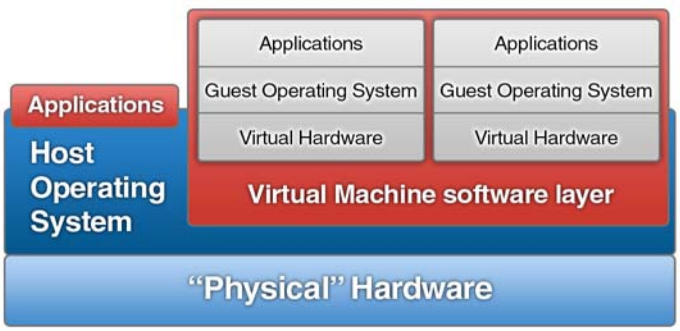

Image9 (Hosted Virtualization)

In summary, server virtualization establishes a symbiotic relationship, so it is important to remember that generating disk I/O in one virtual machine slows I/O to the disk from other virtual systems, no matter the architecture. Fragmentation is both increasingly substantial in virtual machines environments (hierarchical in Hosted Architecture) and compounds the disk bottleneck more so than on conventional systems (shared resource).

For the future, with the opening of proprietary formats for third party development, virtualization-ready hardware from Intel and AMD (improved hardware support and access), operating system advancements (Hypervisor will be an integral part of Windows Longhorn) and technology partnerships such as that between Diskeeper and Microsoft, look for continuing improvements to disk performance as virtualization gets further entrenched in everyday IT.

Improving Disk Performance for Virtual Servers:

1. Run advanced automatic defragmentation on Host and all Guest operating systems. Make sure the software is disk-resource friendly, as intrusive defragmentation with manual or basic tools can cause more problems than they solve.

2. Use basic built-in or more advanced third-party tools (such as Diskeeper's FragShield) to properly adjust MFT and paging file settings on the host and/or guest operating systems.

3. Incorporate the following industry expert recommended disk subsystem and partitioning strategies:

- To minimize fragmentation of the virtual disks, pre-allocate their size if possible – just be sure to monitor and quota disk space.
- Separate the host operating system onto a separate physical disk (spindle).
- Keep the host paging file on a physical disk (spindle) separate from those of the virtual disks.
- Create separate logical partitions on the host system for each virtual system. This is especially true for dynamically expanding virtual hard disks.
- Use high performance SCSI hard disks, and ideally a SAN or RAID back end.
- Compact the virtual hard disks. Make sure to address fragmentation as part of compaction process, both before and after compaction.

Best Practices References:

VM Performance Tips – (Michael Otey - Windows IT Pro)

How to Improve Disk I/O Performances with VMware Workstation 5 - (Alessandro Perilli)

Virtual Server Improving Performance – Microsoft Corporation

Configuring Hard Disk Storage in a Virtual Machine – VMware Inc.

Diskeeper and Microsoft Collaborate to Offer Advanced Virtual Server Performance – Microsoft Corporation

Virtual Server 2005 Best Practices – Microsoft Corporation

VMware Workstation and GSX Performance and Configuration – VMware Inc.

© 2006 Diskeeper Corporation. All Rights Reserved. Diskeeper, The Number One Automatic Defragmenter, and Frag Shield are either registered trademarks or trademarks of Diskeeper Corporation. Microsoft, Windows and Virtual Server are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. VMware is a trademark of VMware, Inc. Virtuozzo is a trademark of SWsoft, Inc. All other trademarks are the property of their respective owners. **Diskeeper Corporation,** 7590 N. Glenoaks Blvd.,Burbank, CA 91504, 800-829-6468, www.diskeeper.com