IDC TECHNOLOGY SPOTLIGHT

Managing the I/O Explosion Without Additional Hardware

December 2012

Adapted from *Worldwide Storage and Virtualized x86 Environments 2012–2016 Forecast* by Ashish Nadkarni, Gary Chen, and Natalya Yezhkova, IDC #235868

Sponsored by Condusiv Technologies

The digital economy has put greater value on information as a competitive advantage. The megatrends of virtualization, cloud computing, and bring-your-own device (BYOD) and the increased use and generation of large digital files are all adding to the mix as enterprises must access, manage, store, analyze, and recover information to increase revenue. This Technology Spotlight discusses the unexpected input/output (I/O) explosion — the direct by-product of virtualization in a digital economy trying to extract value from this compounding data. The document offers guidance to organizations that are looking to eliminate I/O bottlenecks without incurring the cost of extra storage hardware to keep up with application demand. It also examines software optimization technologies such as Condusiv Technologies' V-locity VM, which optimizes I/O at the server level to prevent unnecessary I/O operations from getting into the network or shared storage.

Virtualization — Promise Unrealized

Applications, mobile devices, and embedded systems are creating tremendous amounts of data, with information generated increasing by a factor of 50 and the number of files a datacenter must manage growing by a factor of 75 or more over the next 10 years. To address this data explosion and extract value from it, organizations are implementing virtualization. While virtualization provides enterprises with the agility, scalability, and cost savings required to meet these increasing business and data challenges, it is creating an I/O explosion. Left unanswered, this I/O explosion forces enterprises to purchase ever-increasing amounts of hardware and hire IT staff with expertise in virtualization to manage it all. The combination of hardware and highly skilled people increases investment costs — robbing organizations of the full value of their virtualization investments.

The Information-Based Economy and the Growth of Data

The amount of data generated by the digital universe continues to outpace the capacity to cost-effectively process, analyze, and store the information. The volume of digital data created, stored, and replicated annually now regularly surpasses the zettabyte level (1 sextillion bytes). The vast majority of the information, 75%, is generated by individuals in this age of personal mobile devices. However, enterprises are liable (wholly or in part) for an estimated 80% of this data at some point in its life.

Adding to this load is the BYOD phenomenon, in which mobile workers are using their own smartphones, tablets, and related technologies to generate more information. Because much of this data is unstructured, there must be tools to create data about data — metadata — to help enterprises use the information to their advantage, further adding to the growth and stressing already overwhelmed IT infrastructures.

The biggest driver of this data explosion is financial gain. There is huge economic pressure to extract value from the growing digital universe by processing the right information from the data generated. Moreover, with the recent economic slowdown, enterprises are on a mission to gain more value from information with fewer overall IT expenditures. Virtualization was the first step in accomplishing that goal. However, as much as agility and cost savings were seen at the server level, an unforeseen I/O explosion resulted from virtual machine (VM) sprawl, which has created performance bottlenecks on the shared storage layer. This performance bottleneck issue is crippling many organizations from flexing the full muscle of their virtual environment. As a result, IT administrators are now finding themselves trading cost savings at the server level for increased costs on the storage level to keep up with the new I/O demand in a virtual environment.

While hardware price per performance costs continue to come down, performance improvements are not keeping up with the rate of data growth and the need to extract value from that data. This performance chasm forces enterprises to purchase ever-increasing amounts of hardware. Moreover, enterprises have to hire IT staff with expertise in virtualization and cloud computing to manage all this hardware. The combination of hardware and the hiring of highly skilled personnel increases investment costs. As one customer stated, "When does cheap hardware stop becoming cheap? When I have to buy a lot of it and hire people to manage it."

Although virtualization provides the agility and scalability required to meet increasing data challenges and helps with cost savings, it is aggravating the I/O bottleneck problem. Virtualization enables enterprises to quickly add more VMs as required. The increase in VM density per physical server is rising and with it more I/O, which further constrains network storage. All VMs share the available I/O bandwidth, which doesn't scale well from a cost standpoint with increased processing demands. Unless enterprises address this I/O challenge and its direct relationship to storage hardware, they won't be able to realize the full value of virtualization.

More Value from More Data: The Rise of Virtualization and the Cloud

Because enterprise information is money, executives are looking to technology and IT to create more value faster, using fewer resources. Virtualization has grown because businesses are using it to convert IT into a revenue growth engine that increases corporate value. As a result, virtualization solutions now have built-in intelligent capabilities to link virtual machines with storage and networks. Storage is optimized for virtualized resources, and server virtualization includes mission-critical applications previously off-limits for x86-based virtualization.

Virtualized environments, at the server and desktop levels, will increase and ultimately account for the majority of enterprise data storage capacity. IDC forecasts that by 2015, storage attached to virtualized x86 environments will be 52.2% of external storage systems capacity. Likewise, while cloud computing accounts for less than 2% of IT spending today, IDC estimates that by 2015, nearly 20% of information will be stored or processed in a cloud sometime in its life cycle. IDC believes as much as 10% of all information will be maintained in a cloud.

Migration to the cloud has been driven by the widespread adoption of virtualization. As the use of virtualization and cloud computing increases, the need to access, process, store, manage, and secure the information in the cloud will become more complex as well. For example, over the next 10 years, the number of virtual and physical servers (mostly virtual) will grow tenfold, the amount of information generated will increase by a factor of 50, and the number of files a datacenter must manage will grow by a factor of 75 or more.

As virtualization and cloud computing become further entrenched in the datacenter, they continue to impact other components of IT infrastructure. Storage, in particular, will grow in importance, not only in enabling cloud computing and virtualization but also in gaining value from the vast amounts of information generated by these technologies. Improved performance without additional hardware is required to realize the full value of virtualization and cloud investments.

The I/O Bottleneck

As virtualization and cloud computing have been embraced by the mainstream, an unexpected I/O growth problem has occurred. Storage infrastructures are lagging behind in their cost-efficiency in solving this I/O explosion, which is causing organizations to not realize the full value of their virtualization investments. Because the effectiveness of server virtualization is intrinsically connected to storage performance, I/O performance between server and storage has to be optimized swiftly, automatically, and cost-effectively to help IT deliver data anywhere while meeting service levels, especially in virtual environments. Failing to achieve these objectives will only hamper the ability of IT to support the business in its efforts to cut back costs and grow revenue.

In a survey of enterprise IT executives regarding storage challenges, IDC found that enterprises are feeling the pressures put upon storage architectures. In addition to archiving the vast amounts of data generated by the digital world, respondents listed supporting virtual machines as a challenge.

Many of these challenges are related to I/O performance and capacity scaling, as limitations in infrastructure and an increase in data and metadata are a dangerous combination. In particular, virtual machines — and the growing list of business-critical applications running on them — can be more unpredictable. While enterprises can implement storage area networks (SANs) or build out infrastructure, it's difficult to determine I/O demands in such a fluid environment. For example, if application transaction volume increases, database access from the application server increases as well. Limited network bandwidth in peak times and lower than necessary database performance may also be bottlenecks.

Moreover, the performance gap between the processor and traditional disk-based storage continues to grow. Compute capabilities of current CPUs, especially with multicore technologies, are increasing at a much faster rate than the performance of storage and I/O subsystems based on magnetic media. Server virtualization further exacerbates the performance gap by concentrating application I/O demands across fewer and fewer servers and randomizing these I/O requests with more and more VMs. In an attempt to solve the resulting I/O bottleneck challenge, IT administrators add disk drives and use wide stripe to provide I/O performance. However, this approach increases hardware and management costs.

While IT can increase capacity or implement load balancing, inherent behaviors associated with x86 virtual environments still cause I/O bottlenecks. For example, files written to a general-purpose local disk file system typically are broken into pieces and stored as disparate clusters in the file system. In a virtualized environment, the problem is compounded as virtual machines or virtual disks also can be fragmented, creating even greater need for bandwidth or I/O capacity. Because each piece of a file requires processing cycles — unnecessary I/Os — there can be an increase in overhead, reducing storage and network performance, in effect creating a performance penalty. This extra I/O will slow the speeds of not only a single VM but also other VMs, minimizing the benefits of virtualization.

In addition, virtual disks set to grow dynamically do not shrink when users remove data, creating false requirements for more storage. In environments with many users and large digital files, the bottleneck can grow quickly, dramatically affecting a business.

Considering Condusiv Technologies' V-locity VM Acceleration Software

Burbank, California–based Condusiv Technologies offers high-performance software to optimize I/O to accelerate VM and application performance and efficiency. The company's V-locity VM addresses critical I/O issues by eliminating application bottlenecks in virtualized environments without the need to add storage hardware.

The latest version, V-locity VM offers high-performance storage and application I/O optimization for virtual platforms and in the private cloud. V-locity is a fully certified VMware-ready solution that automatically, intelligently, and transparently improves I/O performance. The software is designed to increase the throughput and efficiency of all VMware ESX/ESXi and Microsoft Hyper-V hypervisor virtual platforms.

V-locity is intended to eliminate unnecessary I/Os at the source caused by the operating system breaking apart files and improves efficiency by optimizing writes for increased bandwidth to the VM. Designed to leverage a SAN or NAS in a virtual environment, the software provides I/O benefit without negatively impacting advanced storage features such as snapshots, replication, data deduplication, and thin provisioning.

In addition, V-locity intelligently caches frequently used files, eliminating unnecessary I/O traffic and improving performance. The software provides faster I/O throughput for VMs and physical server hosts, increased VM density, and optimized VM resource usage within a shared storage system. V-locity also improves storage resource utilization without compromising business productivity.

V-locity VM can be installed in any virtual machine running any of the following versions of Windows: Windows XP, Windows Vista, Windows Server 2003, Windows 7, Windows Server 2008 (or R2), Windows 8, and Windows Server 2012. V-locity supports VMware ESX/ESXi 4.0 and later as well as Microsoft Hyper-V.

V-locity contains a toolkit of technologies to improve I/O performance. IntelliMemory intelligent caching technology boosts frequent file access performance by up to 50% while eliminating unnecessary I/O traffic. IntelliWrite automatically prevents a file system from breaking up a file before it can happen, eliminating all the unnecessary I/Os involved in breaking files into pieces and reading back those pieces. InvisiTasking technology allows all the V-locity background operations within the VM to run with zero resource impact.

Challenges

However, Condusiv Technologies does face some challenges. Given the continued rapid data growth and the increasing role of rich content in the virtual enterprise, the potential for storage and information bottlenecks will only increase in the coming years. Condusiv must continue to enhance V-locity to ensure that the software's capabilities match the exploding data growth.

As virtualization and cloud computing continue to become mainstream enterprise technologies, the problem of I/O bottlenecks will increase, creating more demand for solutions, which will lead to increased competition. Condusiv should prepare for major virtualization and storage players to offer more solutions by maintaining its technological advantage. The company needs to work at changing the market perception that adding more hardware is the way to solve the I/O explosion and its resulting bottlenecks.

Conclusion and Essential Guidance

Virtualization and cloud computing enable enterprises to gain value from the vast amounts of structured and unstructured data generated in today's digital economy. This will increase the dependence on data storage technologies. However, the rise in virtualization and information has

created an unforeseen I/O explosion that dramatically increases the potential for I/O bottlenecks. As such, companies are finding it more difficult to manage and gain value from growing volumes of corporate data in an era of more stringent SLAs, ironically driven by virtualization itself.

The continued expansion of business-critical information and rich content within public and private clouds and content-centric organizations is changing the storage dynamic in enterprises. Extracting value from an expanding universe of digital information is becoming a core business mandate. IDC believes that IT organizations will place a higher priority on boosting the efficiency and reliability of their storage environments for collecting, storing, protecting, and delivering this growing pool of information.

IDC believes that to gain more value from information, enterprises must address the I/O bottleneck issue in the following ways:

Acknowledge that dramatic increases in the amount and use of data have made it difficult to manage storage assets

Determine acceptable performance levels and initiate measures to meet them

Assess multiple storage and networking strategies to determine which will help them gain efficiencies in their virtualized environments

Acknowledge that, due to the nature of virtualized x86 operating environments, I/O bottlenecks will occur, and look for appropriate solutions to mitigate those bottlenecks

To determine the best solutions for eliminating I/O bottlenecks, organizations should seek vendors that have experience not only in I/O optimization but also in providing solutions for virtualized environments. Naturally, solutions should not add to processing overhead and address I/O bottlenecks using multiple approaches. Indeed, solutions that increase VM performance without adding any additional hardware will help enterprises realize the full value of their virtualization and cloud investments.

Information continues to grow not only in volume but also in value. Virtualization and cloud computing provide a way for enterprises to manage information more cheaply and efficiently, but self-restricting challenges are associated with these technologies. It's imperative that businesses realize that these technologies are not a panacea and take steps to mitigate these issues. To the extent that Condusiv Technologies can address the challenges described in this paper, the company has a significant opportunity with enterprises looking to optimize their virtualized IT environments.