Condusiv™ Technologies

WHITE PAPER

How to Accelerate Physical Servers by 50% without Additional Hardware

Executive Summary

There is more data in motion today than ever before. Organizations are no longer focused on data storage as an end; rather they have circled the wagons to focus on extracting value from the massive amount of data they have. To do so, they need a way to capitalize on megatrends like big data, Bring Your Own Device (BYOD), cloud computing and more. They need high-performance solutions that are as effective as they are scalable. Along with the promise of technology megatrends, comes the challenge of managing the I/O chaos that has resulted from them—placing enormous pressure on servers, network and storage. Traditionally, IT has had only one way to confront the I/O problem—throw more hardware in the mix. Condusiv Technologies offers a new approach, both proactive and far more efficient: software that optimizes I/O before it is pushed into the network and storage. The benefit is substantial—performance increases of 50% or more are not only typical, they are expected. I/O optimization software at the top of the stack eliminates unnecessary I/O proactively, before performance penalties are incurred. This paper examines the concept of optimized I/O, and how V-locity performance acceleration software enables IT to efficiently handle the most demanding workloads, support more users, and get the performance they need from hardware they already have.

Megatrends: The Good, The Bad, The Ugly

"Organizations need to tackle the I/O problem in a scalable, cost-effective way."

Big data, BYOD, and cloud computing are megatrends that are equally as promising to business today as they are challenging for IT to manage. Big data is valuable only if organizations can access it quickly and reliably—exactly when it's needed for analytics, reports on-the-fly, and to support critical decision-making. Big data is not worth much if it is not also cost-effective and fast.

These megatrends have created heavier workloads and I/O chaos from a heterogeneous mix of devices and the need for more data in motion that overwhelms the host, network and storage. Organizations trade the benefits of these promising trends for bottlenecks and performance problems, spending more on hardware to compensate for the massive increase in disparate I/O.

Increased Demand for Application and System Performance, Throttled by Increasing I/O

With more data in motion than ever, and greater demand for high performance of that data, organizations need to tackle the I/O problem in a scalable, cost-effective way. Even as organizations move to virtualized environments, I/O-intensive applications like SQL Server and Exchange still reside on physical servers.

While storage density has advanced at 60% CAGR, storage performance has advanced at just 10% CAGR. As a result, reading and writing of data has become a significant bottleneck, as applications running on powerful CPUs must wait for data, forcing organizations to continually buy more storage hardware to spread the load.

This growing performance gap is exacerbated by unnecessary I/O that cannot be resolved with hardware alone.

There are two performance barriers at work: 1) Windows creating unnecessary I/O traffic by splitting files upon write, which also impacts subsequent reads, and 2) frequently accessed data

unnecessarily traveling the full distance between storage and back. These two behaviors create a surplus of I/O that prevents data and applications from performing at peak speeds. These behaviors, once resolved, enable performance increases of 50% or more.

With unnecessary I/O constraining servers and network storage, organizations can't scale from a price/performance perspective. Unless this challenge is addressed by optimizing I/O at the source, stopping unnecessary I/O traffic *before* it's pushed into the network and storage, organizations cannot capitalize on the potential of game-changing megatrends and technologies.

One-to-One Server to Storage

What Changed?

As organizations evolve their focus from storing massive amounts of data to finding ways to extract value from that data, there is a paradigm shift from data-at-rest to data-in-motion. Factor in all the random, chaotic I/O from BYOD and other trends, and organizations face a real challenge in getting the performance they need while managing the I/O problem. Storage hasn't always been a bottleneck. In a traditional one-to-one server to storage relationship, a request for blocks of data is organized and efficient.

However, add multitudes of random I/O traffic from a mass of disassociated data access points and I/O hungry systems, all making similar requests for storage blocks, both random and sequential, to a shared storage system, and the result is an explosion of randomized I/O that affects the network, server and storage. Disparate, random blocks penalize storage performance as it attempts to manage this overwhelming data flow. In addition, the restructuring of data blocks into files takes time in the server, therefore impacting server performance. Because server technology supports multiple tasks and I/O requests simultaneously, data on the I/O bus has changed from efficiently streaming blocks to a mass of random, unnecessary I/Os.

The issue isn't how untidy the landscape has become. The issue is performance degradation caused by unnecessary I/Os being pushed through the entire infrastructure.

Megatrends Impacting I/O Behavior

I/O Bottlenecks

I/O bottlenecks are forcing IT managers to buy more storage, not for the capacity, but to spread I/O demand across a greater number of disk drives. Organizations considering solid-state drives (SSDs) in the array are having difficulty justifying the high cost per performance along with capacity limitations.

Performance Barriers Caused by Unnecessary I/O

When a system services many simultaneous requests from a number of disparate data access points, I/O bottlenecks occur. This problem is compounded by write inefficiencies: files written to a general purpose local disk file system are typically broken into pieces by the Windows OS and stored as disparate clusters in the file system. In today's data-driven environment, there is an even greater need for bandwidth or I/O capacity to sort out where to put incoming data written to the volume. Each piece of a file requires its own processing cycles, resulting in an unwelcome increase in overhead that reduces storage and network performance. A better option is to proactively prevent files from being broken into pieces on the server side, and aggregating these pieces into one sequential file to eliminate unnecessary I/Os and increase overall efficiency of the array.

In addition to write inefficiencies, there is a remarkable amount of common data active in today's enterprise. A massive amount of frequently accessed data travels the full distance to storage and back when it could be made available closer to the application, freeing the infrastructure from unnecessary traffic. Caching hot data in available server memory, combined with the efficiency of I/O optimization at the time files are written, has significant payoff: all writes and subsequent reads become super-efficient, dramatically increasing performance across the entire environment.

"When a system services many simultaneous requests from a number of disparate data access points, I/O bottlenecks occur."

Increasing I/O Workload

After a file leaves the file system, it flows through server initiators (HBAs). As server processor performance grows, organizations tend to increase the workload executed by each server. However, when I/O performance exceeds the HBA's Queue Depth, I/O latency increases, which causes sluggish applications. Avoiding I/O throttling is another reason organizations buy extra HBAs and servers—making additional investments in server hardware to handle growing I/O demand.

Storage Connectivity Limitations

Storage connectivity limitations further impact performance and scalability. When addressing this issue, organizations need to consider the number of hosts they can connect to each array. This number depends on:

"As server processor performance grows, organizations tend to increase the workload, creating bottlenecks."

- The available queues per physical storage port
- The total number of storage ports
- The array's available bandwidth

Storage ports have varying queue depths, from 256 queues to as many as 2,048 per port and beyond. The number of initiators a single storage port can support is directly related to the storage port's available queues. For example, a port with 512 queues and a typical LUN queue depth of 32 can support up to: 512 / 32 = 16 LUNs on 1 Initiator or 16 Initiators with 1 LUN each, or any combination that doesn't exceed this limit. Array configurations that ignore these guidelines are in danger of experiencing QFULL conditions in which the target/storage port is unable to process more I/O requests. When a QFULL occurs, the initiator will throttle I/O to the storage port, which means application response times will increase and I/O activity will decrease. Avoiding I/O throttling is another reason IT administrators buy extra storage to spread I/O.

Storage Farm Inefficiencies

The storage farm presents additional challenges that impact I/O. Rotational disks have built-in physical delays for every I/O operation processed. The total time to access a block of data is the sum of rotational latency, average seek time, transfer time, command overhead, propagation delay and switching time.

At the same time, other factors compound the bottleneck: for example, when users want to write to the disk, write delay penalties will be incurred; blocks coming out of the array will be random and inefficient from a transfer performance point of view; and the storage subsystem will have enough delay that an I/O queue will build, further impacting performance of the storage subsystem and the server.

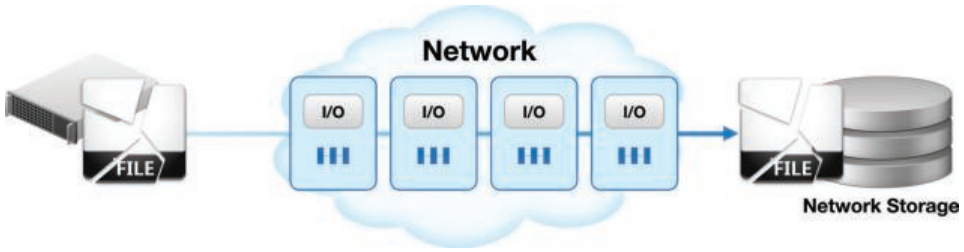
One emerging trend is to use SSDs in the array to eliminate some I/O overhead. However, this is an expensive solution, with limited capacity. A better option is to move frequently accessed files to available server memory, reducing the number of I/Os the storage array must manage. This means the SSDs you have in your environment will now be processing productive I/O instead of spinning unnecessary cycles.

"Failing to overcome the I/O bottleneck will constrain IT's ability to support more users and take advantage of today's megatrends to get more value from their data."

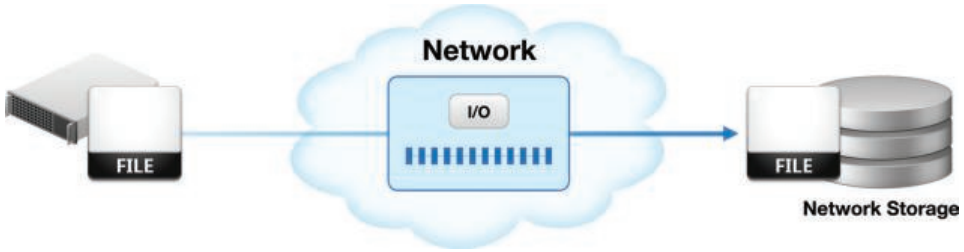
Failing to overcome the I/O bottleneck will constrain IT's ability to support more users and manage increasingly data-centric applications.

Solving the Problem

Before V-locity: Split files created upon write



After V-locity: IntelliWrite writes files contiguously

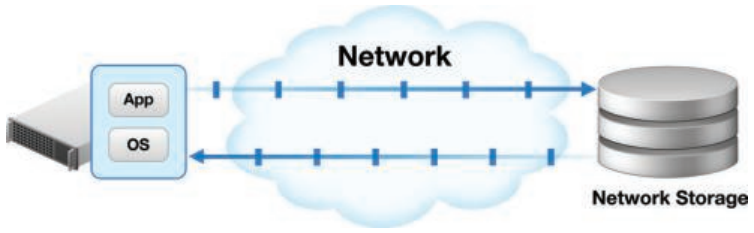


V-locity software optimizes I/O at the source—eliminating unnecessary I/O before it gets pushed into the servers, network and storage. V-locity is composed of two high performance technologies:

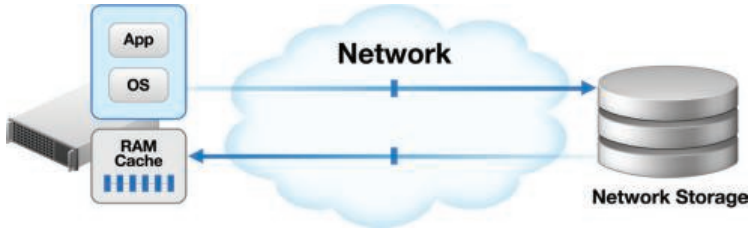
1. V-locity's IntelliWrite® technology eliminates nearly all unnecessary I/O operations at the source when writing a file, which eliminates unnecessary I/O operations on subsequent reads. IntelliWrite aggregates the write so the I/O is fully productive to behave sequentially, increasing throughput by 50% or more.

2. V-locity's IntelliMemory™ caches frequently accessed data within available server memory without memory starvation to the application, since it is dynamically throttled. As a result, read requests no longer travel the full distance to storage and back. IntelliMemory automatically promotes files that are used frequently from disks where they would ordinarily reside, to cache memory inside the server. In this way, IntelliMemory boosts frequent file access performance by up to 50% while eliminating unnecessary I/O traffic.

Before V-locity: Active data path to storage and back



After V-locity: IntelliMemory stores hot data in available server memory



V-locity uses these technologies to intelligently manage disk allocation efficiency and performance. V-locity detects common data that is used frequently enough to warrant performance improvements, then reorganizes random, scattered data into sequential order (or very close to it) as a normal part of an ordinary write. Through "reference of locality," the heads on the disk don't need to be repositioned to request the next block of data, so additional overhead delays are eliminated and performance improves.

"The need to buy additional storage or server hardware has been postponed or eliminated."

Improvements in reference of locality means when data is moved across the I/O bus, there is less overhead and wait time, reducing I/O utilization by 50%. To the end user, overall system performance is seen as improving by 50% or more. The need to buy additional storage or server hardware has been postponed or eliminated.

Moreover, because V-locity optimizes I/O from the top of the stack, it is specifically tailored for environments that leverage a SAN or NAS. Since less I/O traffic is being pushed through the storage, it results in a SAN/NAS that is no longer processing unnecessary I/O requests. SSDs anywhere in the infrastructure will also see performance benefit: V-locity is designed to improve performance from the top all the way down the stack, since the entire infrastructure is now processing only productive I/O.

Conclusion

With V-locity running on physical servers, I/O is optimized at the source. Physical servers perform more work in less time and application latency is dramatically reduced. Optimized I/O means organizations can scale—sharing servers and storage among more devices and more users—all while managing their most demanding workloads and data-centric applications. Since all I/O optimization is taking place at the source, that means V-locity is compatible with all SAN/NAS vendors. And all advanced features on network storage benefit.

More Work, Less Time

openBench Labs recently tested V-locity Server on servers running Exchange 2010 and Active Directory. A 1-hour LoadGen test simulated 1,000 Outlook users and revealed that V-locity Server enabled 75% more transactions in the same amount of time. Latency was dramatically reduced, and the LoadGen test showed a 240% increase in IOPS.

For organizations that manage a mix of physical and virtual servers in their data centers, which run a variety of applications and provide services to an increasingly diverse mix of devices, this means more transactions are processed without lag, enabling much more work in the same amount of time. As applications become increasingly data-centric and systems more I/O hungry, the threat of mounting, chaotic I/O is eliminated, *before* it can ever inhibit performance and impact budget.

"I/O optimization means business can evolve, capitalizing on the promise of significant technology trends."

I/O optimization means business can evolve, capitalizing on the promise of significant technology trends. With V-locity Server, organizations can drive 50% or more performance from the hardware they already have. Not only does this protect the investment of their infrastructure, it allows them to be more competitive—to invest in strategic initiatives; not in unnecessary hardware.

Think Faster.™ Visit us at **Condusiv.com**

Condusiv Technologies
7590 N. Glenoaks Blvd.
Burbank, CA 91504
800-829-6468
www.condusiv.com