The Server-Storage Performance Gap

How disk drive throughput and access time affect performance

Gear6 accelerates storage and delivers real time performance. Deployed in the data center, Gear6 products and solutions provide scalable and transparent acceleration for existing storage infrastructures to speed up applications, eliminate peak load disruptions, and simplify overall management.

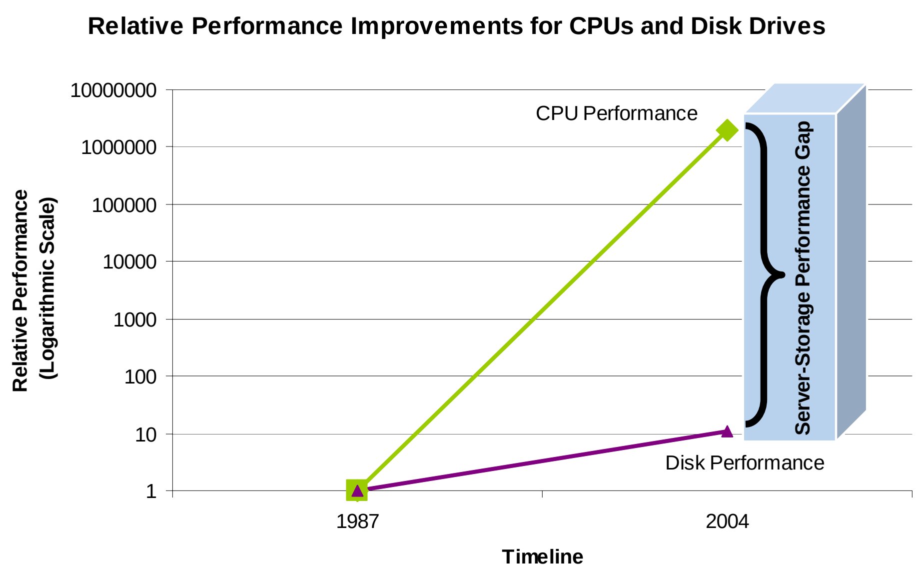
INTRODUCTION

In enterprise storage configurations and data centers, hard disk drives serve as the foundation of information operations.

While we are bound to have disk drives for a long time, current disk-based storage cannot keep up with the increasing processing capabilities of powerful servers and appetite of data-intensive applications. The existing disk storage infrastructure needs assistance, and it is time to properly apply advanced technology to solve this performance gap.

Historically, different computer system components have advanced performance at different relative rates. Although disk capacity has improved somewhat, disk performance ranks at the bottom with no significant improvement compared to million-fold boosts by other system components. For example, CPU performance has progressed at an impressive clip, driven by Moore's law, multi-core processors, and threading technology to increase 2,000,000 times since 1987. In comparison, disk performance only improved by 11 times. This has created a significant and growing Server-Storage Performance Gap shown in Figure 1. Note the multiple orders of magnitude difference between CPU and disk increases made visible by a logarithmic scale.

Figure 1
Relative Performance Improvement of System Components
Source: Seagate, see appendix



The impact of the Server-Storage Performance Gap has led to a wide variety of methods to boost data throughput and data access performance. These methods are so entrenched in storage and data center thinking that we almost take them as conventional wisdom. Table 1 outlines several examples and their intent.

Table 1
Traditional methods to increase disk performance

Method	Intent
RAID	Stripe data across multiple disks to increase the overall bandwidth of the storage system
Short stroking	Place data only on outer section of disk platters to reduce the seek time of the mechanical heads
Thin Provisioning	Use virtual volumes to constrain data to small areas of a physical disk
Parallel file systems	Link hundreds or thousands of individual disk spindles to increase system throughput
Virtualization	Similar to parallel file systems, use software to spread volumes across multiple drives; a more sophisticated extension of RAID 0
Faster drives	Choose 15K RPM drives over 10K RPM drives to increase speed

All of the above methods are ultimately constrained by physical disk latency and require deploying additional mechanical infrastructure for improvement. This leads to fallout requirements in space, power, and cooling.

PERFORMANCE CHARACTERISTICS AND INHIBITORS

To better understand the underlying inhibitors for disk drive performance in enterprise computing systems, we need to categorize three primary metrics relevant to architecting disk-based storage systems: capacity, throughput, and access time.

Capacity

Capacity remains relatively unconstrained compared to other metrics, and in general is not considered a performance metric. That said, many IT and storage managers end up with excess capacity as an indirect result of performance tuning, requiring more valuable space, expensive power, and cooling.

The more important questions to ask are 1) how much data can I access simultaneously (throughput), and 2) how long does it take each time I want to retrieve a piece of data (access time).

Throughput

Throughput is the first of two disk performance metrics, and represents the amount of data that a disk drive, or a group of disk drives, can deliver at any given moment. For disk drives, this term is often referred to as I/O Operations Per Second (IOPs) which usually ranges in the low hundreds for a single disk.

Since applications typically require far greater IOPs that what a single disk can deliver, multiple disks are linked together to form a subsystem. Adding more disks to the subsystem increases the overall throughput. But there are significant downsides to this approach in that adding more disks just to get higher bandwidth means an increase in power, space, and cooling requirements. It also means taking care of more mechanical parts that have a higher failure rate compared to non-mechanical components. Plus there is the cost of buying and supporting additional unnecessary spindles.

Virtualization at the disk or subsystem level, and parallel file systems are both examples of using more spindles to increase bandwidth.

Throughput helps with applications that require a lot of information simultaneously in areas such as:

- Media, content, and animation
- Energy and Exploration
- Biotechnology
- Electronic Design Automation

Access Time

Access time is the second disk performance metric and represents the length of time it takes for a disk to seek, find, and deliver data. With mechanical disks, it takes time to move the actuator, or read head, to reach the specific disk platter location, and then the disk can only rotate at a certain number of Revolutions Per Minute (RPM).

Access time directly impacts the performance of applications and databases that handle numerous time sensitive transactions such as:

- On Line Transaction Processing (OLTP)
- Business Intelligence, On Line Analytical Processing (OLAP)
- Telecommunications
- Financial Services

In these scenarios, requests accumulate in a queue. During peak loads, each request must wait for the preceding operations to complete. So the total time for a request to complete is the data access time for that individual request, plus a time multiplier for all of the preceding operations. When an individual request takes milliseconds, the overall impact on the queue becomes significant.

TACKLING THE GAP

IT and storage managers attempt to tackle the Gap in a number of ways such as using RAID, short stroking, parallel file systems, virtualization, and other schemes that seek to overcome the limitations of mechanical disks.

While these means can increase throughput, they can create costly and complex overhead by requiring management of an ever-increasing number of spindles.

More importantly, adding spindles can not accomplish the pressing need to reduce data access time. By reducing access time, servers can reach more data more quickly, avoiding data request queues. This allows a greater number of users to share a single data set, eliminates disruption from peak loads, and reduces management cost and complexity.

Tackling the Server-Storage Performance Gap requires consideration of all disk drive characteristics. The winning combination should work with existing capacity, boost throughput, but also reduce access time, while relying on simpler configurations.

Appendix

Data for the Server-Storage Performance Gap chart

	1987	2004	Increase
CPU Performance	1 MIPS	2,000,000 MIPS	2,000,000x
Memory Size	16 Kbytes	32 Gbytes	2,000,000x
Memory Performance	100 usec	2 nsec	50,000x
Disk Drive Capacity	20 Mbytes	300 Gbytes	15,000x
Disk Drive Performance	60 msec	5.3 msec	11x

Source: Seagate Whitepaper TP-525,
http://www.seagate.com/docs/pdf/whitepaper/economies_capacity_spd_tp.pdf

Glossary

CPU – Central Processing Unit

Giga Byte (GB) – Billion Bytes

IOPs – Input/Output Operations per second

Kilo Byte (KB) - Thousand Bytes

Mega Byte (MB) - Million Bytes

Microsecond (us) – One millionth of a second

Millisecond (ms) – One thousandth of a second

MIPS – Millions of Instructions Per Second

Nansosecond (ns) – One billionth of a second

RPM – Revolutions Per Minute

gear6