

- internet.com
- IT
- Developer
- News
- Small Business
- Personal Tech
- Freelance Jobs **new!**
- Partners
- Solutions
- Shop
- Login
- Register

close X

CLICK HERE

privacy policy

- IT MANAGEMENT
- NETWORKING
- WEB DEVELOPMENT
- HARDWARE & SYSTEMS
- SOFTWARE DEVELOPMENT
- IT NEWS



- Events
- Jobs
- Premium Services
- Media Kit
- Network Map
- E-mail Offers
- Vendor Solutions
- Webcasts

subjects:

- Business Continuity +
- Hardware +
- Industry News
- IP Storage +
- Management +
- Outsourcing +
- SANs/NAS +
- Technology +
- Special Reports
- Videos
- Products
- Glossary
- Events
- Technology Jobs
- EnterpriseStorageForum Webcasts:
- Storage Is Changing Fast - Be Ready or Be Left Behind

more Webcasts...

Search EarthWeb Network

Find a Storage Term



HP PartnerONE | SolutionsINFINITE Visit us at hp.com/partners/us/go/4

[enterprisestorageforum.com](#) : [Technology](#) : [Technology Features](#): [I/O Bottlenecks: Biggest Threat to Data Storage](#)

I/O Bottlenecks: Biggest Threat to Data Storage

December 31, 2009

By [Henry Newman](#)



Data storage has become the weak link in enterprise applications, and without a concerted effort on the part of storage vendors, the technology is in danger of becoming irrelevant. The [I/O](#) stack just isn't keeping pace with advances in silicon, and it could find itself replaced by new technologies like [phase change memory](#) (PCM) that promise unfettered access to data.

The problem is simple: Memory bandwidth and CPU performance continue to grow much faster than disk and bus performance and disk channel speed, and combined with a limited I/O interface ([POSIX](#)), the result is in an I/O bottleneck that only gets worse with time.

A look at the performance increases for various elements of the storage stack over the last five years paints a clear picture:

Microsoft: IT Projects Unified Communications

Improve Communications & Collaboration
Intel expects 20 percent reduction in audio conferencing costs with Unified Communications. Learn more.

[Click Here](#)

[Employee Internet Monitoring Tools](#)

[TimeClock Plus - Employee Time & Attendance Software](#)

[Network Performance Monitoring Software](#)

[Document Management Services](#)

[Employee Internet Monitoring Tools](#)

[Autodesk Inventor: Revolutionizing Digital Prototyping](#)

[Recruiting and Staffing Software](#)

[Contact Management Software - Hoover's®](#)

- **Memory bandwidth:** Intel has gone from 4.3 GB/sec in 2004 to 40 GB/sec, while AMD has gone from 5.3 to 25.6, an increase of 9.3 times for Intel and 4.8 times for AMD.
- **CPU performance:** Using Moore's Law that transistor count doubles every 18 months, I will assume that this translates to performance (which it does not) for a greater than tenfold performance improvement.
- **PCIe bus** has increased from 250 MB/sec per lane in 2004 to 500 MB/sec per lane in 2008 and an expected 1 GB/sec next year, an increase of two (and soon four) times.
- **Per disk channel speed** has increased by 50 percent, from 4Gb Fibre Channel to Gb SAS.
- **Disk performance** for SATA has improved from 68 MB/sec to 84 MB/sec, while FC/SAS has gone from 125 MB/sec to 194 MB/sec, a modest improvement similar to per disk channel speed in the case of FC and SAS, while SATA has improved only about 24 percent.
- **Disk density** has doubled for FC/SAS, from 300 GB to 600 GB, while SATA has seen an eightfold increase from 250 GB to 2 TB.

As you can see, disk channel speed and performance and the [PCIe](#) bus have lagged CPU and memory bandwidth performance. What this means is that access to data can't keep up with the ability to process it, and POSIX is unable to determine what's important or define [quality of service](#), problems that could spell doom for storage networking as other options become available.

MARKETPLACE

[Windows 7 will benefit your business](#)

Learn how the new features in Windows 7 will improve your IT and Ops efficiency.

www.microsoft.com

[Free Guide to IT Asset Management](#)

Free 10 page Buyer's Guide for IT with Detailed Feature Checklist. Read It Now.

KACE.com/IT-Guide

[Microsoft Exchange Server 2010](#)

Simplify Administration and Deployment of Messaging - Free Download.

www.microsoft.com

[HP Notebook Smart Choice Super 9](#)

Click Here to Find Out More About Great Rebates & FREE Access W/Bundles!

www.synnex.com

[Advertise](#)

Related Articles

- [The Future of Data Storage: FCoE, SSD Mergers, No Clouds](#)
- [Top Storage Stories of 2009: RAID, Clouds, SSDs and Mergers](#)
- [Error Correction: An Urgent Need for Files](#)
- [Seagate Unveils First Solid State Drives](#)
- [AMD Does RAID On a Chip](#)

Enterprise Storage Glossary

[Fibre Channel](#)
[ILM](#)

internet.commerce

[Be a Commerce Partner](#)

[Phone Cards](#)

[IP Services](#)

[Car Donations](#)

[Website Hosting](#)

[Business Liability](#)

[Dedicated Servers](#)

[Premium Bandwidth](#)

[PDA Phones & Cases](#)

[Online Universities](#)

[Online Education](#)

[IT Legal Contracts](#)
[Managed Web Hosting](#)
[Product Reviews](#)
[Memory Upgrades](#)

Log Analysis Platform

Log analysis reports for application, network and security log, download now!
www.xpolog.com

Access to data is not increasing anywhere in the system, not the PCIe bus, not the storage channel and not the disk drive. The only exception is [solid state drives](#), and specifically [SLC SSDs](#). There was nothing to compare these drives with five years ago, but they are nonetheless still limited by the channel speed, which is currently at 6Gb SAS, or 768 MB/sec for a channel.

[iSCSI](#)
[JBOD](#)
[NAS](#)
[RAID](#)
[SAN](#)
[SAS](#)
[SATA](#)
[Virtualization](#)

Search for more storage terms ...

Go!

XML/RSS feeds

[XML](#) [RSS](#)

[EarthWeb IT Management news and headlines](#)

[EnterpriseStorageForum Headlines](#)

[See more EarthWeb Network RSS feeds](#)

FREE Tech Newsletters

Enterprise Networking Planet

[Sign Up](#)

POSIX: Part of the Problem

One problem I see is that we have a minimal interface between applications and storage. All we have is standard POSIX control, with open, read, write and close and a few other system calls. There

is no method to provide hints about file usage; for example, you might want to have a that says the file will be read sequentially, or a hint that a file might be over written. There are lots of possible hints, but there is no standard way of providing file hints, although some file systems have non-standard methods, but the interface is only part of the problem. The real problem is that storage performance is falling further and further below the performance of the system, making storage harder to use and potentially spelling its demise for all but essential usage. I think that there are lots of reasons for the performance lag, but here are a few:

- Designing PCIe bus interfaces to memory is difficult and expensive and having too many PCIe buses connected to memory increases the system cost dramatically.
- Since PCIe bus performance is not keeping pace with CPU and memory advances why do we need faster and faster channel performance?
- Disk drive interfaces are using the same technology as the channel, so the performance is not going to improve much.

This problem has troubled me for a long time. I continually ask myself why the situation hasn't changed, and I keep coming back to the same reason: It is costly and difficult to increase the performance in the storage stack. A change here or there will not change the performance of the whole data path. Simple bottleneck analysis proves this, as changing all but one component means there is still a bottleneck. The whole stack needs to change to improve performance, from the PCIe bus to the device and everything in between, including the software interface. No one company owns the stack. PCIe is a standard, and SAS interfaces and performance is a standard. No one is going to build something that is non-standard, as the engineering costs are just too high and the market is too small.

There are a number of companies building PCIe-based flash storage devices, which eliminate the channel interface problem and some of the performance problems. At least right now, the cost of these devices is far too high for the consumer market, which like or not drives the storage market. Even with that, the performance bottleneck is the PC bus, which has not improved that much and likely will not improve that much in the future. What is needed is something at home that demands more performance.

The one technology that requires bandwidth that far exceeds the PCIe bus is SSDs. The current crop of SSDs can run at over 500 MB/sec for read, which is about 65 percent of single SAS channel; for a SATA drive it would be a peak of about 105 MB/sec, or about 25 percent of the channel, and for SAS drives a peak of 194 MB/sec or about 25 percent of the channel. The bottom line is it does not take too many disk drives to saturate a SAS channel today. Early in the Fibre Channel days, in 1998, a two FC loop was capable of supporting the I/O for nine disk drives. Today two 6Gbit SAS channels with enterprise drives will support a tad fewer number of SAS drives, so there has been no improvement. And we just got to 6Gbit SAS technology and most RAID systems are still using 3Gbit, which means that you can support about 1.9 drives of today's fastest SAS drives.

Phase Change Memory Could Doom Storage Networks

If vendors cannot find a way to fix the I/O stack — which would involve a lot of groups, such as OpenGroup, ANSI T10, T11 and T13 and PCI SIG — I think the I/O stack might be the way of eight track tapes and floppy disks. Given the difficulty of solving the hardware and interface problem, the situation appears ripe for someone to figure out a workaround. The most obvious solution is to do as little I/O as possible. Right now there is a limitation on the number of memory channels that are available on commodity CPUs and there is also a limit on performance and distance from the CPU for DIMMs. Since NAND flash is byte addressable, it makes its potential as a memory extension more difficult, but phase change memory (PCM) does not have this limitation. For a good overview of PCM, see a presentation from Numonyx, which is working on PCM technology with Intel (NASDAQ: INTC).

“PCM has the potential to move data storage ... from the center of data centers to the periphery. ... In short, disk becomes the new tape.”

While years away, PCM has the potential to move data storage and storage networks from the center of data centers to the periphery. I/O would only have to be conducted at the start and end of the day, with data parked in memory while applications are running. In short, disk becomes the new tape.

The I/O stack can be fixed, but there is no immediate financial incentive to fix it now. Enterprise storage vendors are smart enough to see that their entire industry could be threatened a few years down the road by a new technology like PCM, perhaps they can get to work on fixing the problem now, because it will take time.

The ability to put flash and PCM on the motherboard like DIMMs is coming soon. This may not happen in the next few years, but by the end of the new decade, watch out.

Henry Newman, CTO of [Instrumental Inc.](#) and a regular Enterprise Storage Forum contributor, is an industry consultant with 28 years experience in high-performance computing and storage.

[See more articles by Henry Newman.](#)

Follow Enterprise Storage Forum on [Twitter](#)

Tools:  Email  Print  Digg This Story  del.icio.us

Add www.enterprisestorageforum.com to your favorites

 MY YAHOO!  Windows Favorites

Add www.enterprisestorageforum.com to your browser search box

[IE 7](#) | [Firefox 2.0](#) | [Firefox 1.5.x](#)Receive news via our XML/RSS feed [XML](#) [RSS](#)[Technology Features Archives](#)2 Comments ([click to add your comment](#))By **Jim** January 1 2010 5:08 AMPDT

Interesting article but I think that the basis is incorrect. You talk about 'memory', 'cpu', and 'PCIe' as if they are single entities but in fact over the past few years you have seen the introduction of large numbers of multi-bank meory, multi-core CPUs and multi-lane PCIe. Individually these items have not increased significantly in speed/throughput/performance, people have just learned how to package them in such a way that they work well together.

Disk is doing much the same. Products like Xiotech's ISE and Atrato's Velocity are providing the equivalent of a next-generation disk, with multiple individual units co-operating to provide the performance and capacity benefits without the downsides of traditional arrays. If you compare the individual disk to these products then storage is holding its own quite nicely against the increases in other areas of computer infrastructure.

[Reply to this comment](#)By **Chris Bruner** January 1 2010 7:41 AMPDT

Linux Distros like puppy Linux are already doing what you are talking about, by running in memory. That the compressed disk is loaded into memory and the programs are decompressed as needed. At the end of the session, the user data is saved to disk.

While I don't use puppy myself, (I like big fat user interfaces) the few times that I have used it I have always been impressed by the speed. Which in my mind proves your point.

[Reply to this comment](#)

Comment and Contribute

 Your name/nickname Your email**XHTML:** You can use these tags: `` `<u>` `<i>`(Maximum characters: 1200). You have **1200** characters left.



Please type the alphanumeric characters above and click "Submit" to continue. [What's this?](#)

I cannot read this. Please generate a

[See our comment policy.](#)

STORAGE SOLUTIONS

- ▶ Intel Tech Brief: Establishing a Low Cost Disaster Recovery Site
- ▶ Maximize your business with Intel® Ethernet optimization for Windows Server® 2008 R2
- ▶ Tech Brief: HP BladeSystem and Managed Services Help Save Six Figures Annually
- ▶ Intel Article: The Problem of Power Consumption in Servers
- ▶ Intel Tech Brief: Automated Energy Efficiency for the Intelligent Enterprise



The Network for Technology Professionals

Search:

[About Internet.com](#)

Copyright 2010 QuinStreet Inc. All Rights Reserved.

[Legal Notices](#), [Licensing](#), [Permissions](#), [Privacy Policy](#)

[Advertise](#) | [Newsletters](#) | [E-mail Offers](#)

Solutions

Whitepapers and eBooks

Microsoft Personalized Whitepapers and Resources for You
Microsoft SQL Server 2008 Index
Articles: Build a More Agile Business with IBM
IBM Articles: Virtualization Delivers a Dynamic Infrastructure
Articles: IBM Offers Enhanced Measurement and Management for Energy Usage
Articles: IBM Helps Transformation to an Information-Based Enterprise
Microsoft Articles: Visual Studio 2010

Adobe Article: Adobe Helps PHP Developers Create Rich Internet Applications
Microsoft SharePoint Index
Adobe Article: Java Developers Finding a Home at Adobe Flex
IBM Articles: A Smarter Business Needs Smarter Technology
Avaya Article: Communication-Enabled Mashups--Empowering Both Business Owners and IT
Microsoft Article: Stress Free and Efficient Designer/Developer Collaboration
MORE WHITEPAPERS, EBOOKS, AND ARTICLES

Webcasts

Microsoft Forefront Index
Microsoft Unified Communications Index
Webcast: Connecting PHP to Microsoft Technologies

Video: Windows Azure Debugging Tips
SAP BusinessObjects Webcast: Unlock the Power of Reporting with Crystal Reports
MORE WEBCASTS, PODCASTS, AND VIDEOS

Downloads and eKits

Download: Crystal Reports 2008 Free Trial
Microsoft Windows Server 2008 R2 Index
Downloads & Free Trials: Microsoft's New Efficiency Resource Center
Get the Windows 7 SDK

Free Trial: Red Gate SQL Backup Pro 6
Iron Speed Designer Application Generator
MORE DOWNLOADS, EKITS, AND FREE TRIALS

Tutorials and Demos

Microsoft Article: Boost Your Virtualisation ROI
Internet.com Hot List: Get the Inside Scoop on IT and Developer Products
Microsoft Article: Forefront Offers One-Stop Security Shopping for the Enterprise
All About Botnets

Microsoft Article: Improved Web Collaboration on Tap for 2010
Microsoft System Center Index
MORE TUTORIALS, DEMOS AND STEP-BY-STEP GUIDES