

A Snapshot of the Current State of Linux Storage

by Graeme Thickins
July 2002



TidalWire

The Storage Connection.SM

www.tidalwire.com

You've heard it before: Linux is fine for certain applications, like web hosting or email, but it's not really up to snuff yet for mission-critical data center duty.

Recent developments, however, would seem to suggest that this widely held opinion about Linux could now be, if you'll pardon the expression, "Old Hat." Many think this best known of open-source OSes -- the leading vendor of which is distributor/packager Red Hat Software -- could in fact be ready to play a more serious role in enterprise IT soon, if it isn't in fact already for some firms.

Consider IBM's announcement in June that it's spending \$1 million in a facility in its Manhattan offices just to demonstrate Linux to Wall Street - smack in the middle of a longtime Sun stronghold. It doesn't get any more mission-critical than the world's leading financial nerve center. In fact, according to a recent Business 2.0 article, Merrill Lynch and Credit Suisse First Boston have in fact already begun using Linux throughout their operations, for such high-powered tasks as financial trading and order processing.

But buzzwords often get in the way. "Let's face it, 'mission-critical' means different things to different people," said Gary Doan, CEO of Intradyn, an admitted "stealth-mode" open-source startup in Minneapolis. The meaning of the term can vary with the size of the enterprise, the industry type, and other factors, he said. "It's in the ear of the beholder. Meantime, Linux marches on." Doan will only say his company is broadly focused on private-network and NAS solutions based on Linux.

But more than startups have been active in Linux storage in recent months. The list of vendors making open-source related product announcements reads like an industry-conference attendee list - including Sun itself, as well as Network Appliance, IBM, Oracle, StorageTek, Legato, Quantum, Dell, and more. Several well-positioned emerging firms, like FalconStor, Dot Hill, and Sistina Software, are placing big bets in the Linux storage arena as well.

According to IDC, 28 per cent of current Unix storage is made up of SANs, whereas Linux has only achieved 8.6 per cent penetration so far. But IDC predicts that the value of SANs deployed on Linux will grow from \$58 million in 2001 to \$418 million by 2005 - more than a 6X growth rate.

Others have reported that, although today's overall SAN market is relatively flat, Linux is one sector showing decent growth. One such firm is JNI, a San Diego designer and supplier of enterprise connectivity products, including fibre channel hardware and software.

JNI recently announced two drivers for its 2Gb fibre channel host bus adapters that now support 133-MHz PCI-X device attachments. They're designed to run under Red Hat Advanced Server 2.0 and Red Hat Linux version 7.

Though 66-MHz PCI is still used by most companies, "it's a bottleneck," said Hendrik Wacker, a sales and marketing executive at JNI. "With 133-MHz PCI-X, you get 1.8Gb throughput - the maximum achievable in practice - compared to only 1.1Gb to 1.2Gb with PCI."

Wacker said that Brocade's 3800 FC switch, released earlier this year, has enabled widespread 1Gb to 2Gb upgrade of equipment that supported 2Gb, and that this has triggered demand for PCI-X.



Build It and They Will Come

The current iteration of Linux - the 2.4 kernel - "was the first the enterprise helped build," according to an eWeek Labs article when the kernel was released in early 2001. Its development was driven by the demands of enterprise customers, enterprise applications, and intensive benchmarking - "resulting in the 2.4 kernel having far more of an enterprise focus than ever before," eWeek said at the time. In particular, storage manageability, capacity increases, and SMP (symmetric multiprocessing) were the key areas, they noted, where this kernel took dramatic strides beyond its 2.2 predecessor.

"The 2.4 kernel was a big step forward in storage management," said eWeek Labs, "by providing an LVM (logical volume manager) subsystem, similar to the LVM included in Hewlett-Packard's HP-UX." This code, developed by Sistina Software, let IT managers dynamically add disks to and delete disks from Linux storage pools without rebooting (also a feature in Windows 2000), as well as create file system snapshots for backup.

The Kernel Marches On

Matthew O'Keefe, Ph.D., CTO at Sistina Software, Minneapolis, had this to say about the current state of affairs in Linux and storage: "The good news is that support from the major hardware and storage vendors like EMC, IBM, HP, Veritas, and Brocade is there for Linux. You can get certified hardware and software now for the major Linux distributions. The not-so-good news is that the Linux kernel still needs work to support a larger number of storage devices, and devices larger than two terabytes."

Currently, he noted, Linux is limited to two terabytes per device and, therefore, the same per file system. "There is ongoing work to fix this problem in the Linux 2.5 development kernel," O'Keefe said, speaking as one who's been involved in Linux for several years. His background includes developing and demonstrating the industry's first cluster file system running over a fibre channel SAN at the 1997 NAB conference. He has also published more than 60 technical papers, and is a member of the IEEE Technical Committee on Mass Storage Systems.

"Fundamentally, Linux needs storage area networking to scale to compete with high-end server hardware," he added. "There is no equivalent to an E10000 in terms of high-end PC server hardware -- at least not yet. Therefore, clusters of Linux machines are required to provide high performance, and the only way high I/O throughput and simplified management can be achieved in this environment is with shared SAN storage."

He noted that a cluster file system and volume manager, like Sistina's GFS and CLVM solutions, enable this kind of shared-storage clustering in Linux.

When asked what trends he sees in the evolution of Linux, O'Keefe said: "Larger memory systems, larger storage subsystems, more use of storage area networks, cluster file systems, and cluster volume managers."

And, in a contrarian note, at least compared to some Linux touters, O'Keefe had this to say about how it will continue to be just part of the mix going forward: "Because Linux cannot yet handle the heavy mission critical data center workloads, it's important to integrate Linux storage with Solaris, HP-UX, and AIX storage, because customers are going to be using both kinds of Unix for a *long* time. Linux is great at the low-end and mid-range applications -- web serving, email, data warehousing - but you still need big Unix iron for the heaviest data center jobs."



Mother of All Convergences?

"SAN and NAS convergence is very important," CTO O'Keefe said, "and Sistina's product architecture -- a cluster file system that allows multiple cluster nodes to server NAS storage from a single pool out to many clients -- is critical to making it happen.

"SAN enables NAS by providing a single unified pool of storage to manage, while Sistina's GFS allows multiple NAS heads to serve up the same shared storage. So, you get both simplified management and high performance -- something that has not been possible before with NAS.

"We have a lot of customer interest here," he said, "from both embedded customers and end users who are using NAS clustering to simplify management and increase performance and scalability."

Another perspective on SAN/NAS convergence comes from Geoff Barrall, CTO of BlueArc, a high-profile storage startup based in San Jose (though one that uses its own OS, not Linux). Barrall pointed out in a recent commentary posted on his company's site that none of the major storage vendors "has any doubt that the convergence of SAN and NAS will occur."

Barrall went on to say: "In a market as competitive as the storage market is today, I think we can soon count on storage features being integrated as a single cohesive solution from a single vendor. NAS and SAN convergence is a certainty. But its form, whether network-centric, storage array-centric, or file-centric (NAS), still remains to be decided."

Scientists With Big Storage Needs Push the Envelope

What additional predictions does Sistina's O'Keefe have for Linux storage? "Another trend we're seeing is Linux clusters with SANs and cluster file systems replacing high-end Unix iron when the application does not require high parallel SMP performance, but requires high I/O throughput and lots of storage."

"This is a big trend," he said. "Nearly every high-performance computing customer we're working with is considering this approach, because the price/performance is usually something like 5X better than the old approach." One such customer is Fermilab, which crafted a solution combining Sistina's GFS with Dot Hill's SANnet 4000, for a five-year project in which the lab's astrophysicists will produce a 3D map of the universe.

"The approach many scientific customers are taking involves replacing high-end Sun or SGI SMP with 8-32 processors with a Linux cluster, plus SAN, plus Sistina's GFS cluster file system. The ROI happens almost immediately. Many of these customers are taking hardware maintenance dollars they'd spend for just one year and instead applying it to buying a Linux cluster. It's a better option, they find, than continuing to support an old Sun or SGI SMP machine."

Recent announcements from Sistina have included comprehensive support for Oracle9I, and, earlier in the year, a partnership with FalconStor. In the latter, the two companies will serve customers with lower cost, more flexible IP-based SAN solutions. Together, FalconStor's IPStor™ software suite and Sistina's Global File System™ (GFS) software, the companies say, "will provide large and medium enterprises with the autonomy to create highly secure and available data storage environments, independent of vendor, platform, hardware, or protocol."



Since this announcement, Sistina's O'Keefe noted, "we're working with FalconStor to integrate and test how our products work together. We're pursuing joint opportunities, and have found most interest in the area of remote replication and file sharing."

Wide Open Future?

What lies ahead for storage software vendors, whether Linux or otherwise? Bill North, research director for storage software at IDC, puts it optimistically in his recent forecast report: "Software suppliers whose products simplify and automate the management of storage systems, and those who can keep pace with technology changes and emerging storage architectures, will have rich new opportunities and will continue to experience rapid growth," North said.

Henry Baltazar, senior analyst at eWeek Labs, offers up this perspective: "Open-source is a software-only phenomenon, while good hardware gets patented where possible to preserve value. Beyond the proprietary hardware level, however, open-source technologies are being used to provide added functionality and manageability to appliance solutions."

"Several vendors," he said, "have built advanced storage virtualization features on top of the Linux kernel to create storage appliances that carve up SAN resources."

"We expect to see more storage management appliances flooding the market," analyst Baltazar says, "and we will be very surprised if many of them didn't rely on open-source software."

Links:

Red Hat Software: <http://www.redhat.com>

IBM Linux Portal: <http://www-1.ibm.com/linux/index.shtml>

Falconstor Software: <http://www.falconstor.com>

Dot Hill Systems Corp: <http://www.dothill.com>

Sistina Software: <http://www.sistina.com>

JNI Corporation: <http://www.jni.com>

ABOUT THE AUTHOR

Graeme Thickins is a freelance technology writer and consultant based in the Twin Cities and Southern California. Graeme began working in the storage industry more than 25 years ago and today, among other things, provides news and perspective on storage industry events nationwide. His reporting and commentary has appeared in a number of print and online journals, magazines and newsletters, including [Conferenza](#), [Business 2.0](#), [Office.com](#), [The Journal of the Hyperlinked Organization \(JOHO\)](#), and [NanoTechNews.com](#).

