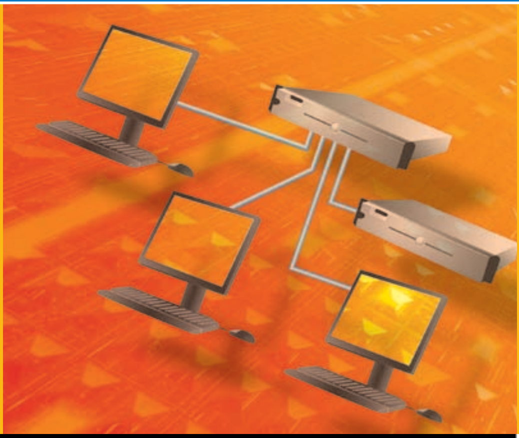

EXPLORING iSCSI AND iSCSI BOOT FOR SAN IMPLEMENTATIONS

BY ROBERT LUSINSKY
DAVID CHRISTENSEN

The ability to boot server operating systems over an Internet SCSI (iSCSI)–based storage area network can offer multiple advantages, including increased system reliability, simplified management, and accelerated restore processes. To help enterprises utilize this functionality, Broadcom and Dell have teamed up to offer enhanced iSCSI boot on Broadcom® Ethernet controllers in Dell™ PowerEdge™ servers.

Related Categories:

Broadcom

Internet SCSI (iSCSI)

Storage

Storage area network (SAN)

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

The Internet SCSI (iSCSI) protocol provides a method of transporting traditional SCSI commands, data, and status messages over standard TCP/IP networks. iSCSI offers a simple, cost-effective, standards-based way for even enterprises with modest budgets to deploy networked storage and gain the advantages of block-level storage consolidation. And because the familiar SCSI interface is maintained over standard networking topologies, iSCSI can help simplify networked storage implementation for both storage and network administrators.

iSCSI boot allows administrators to boot server operating systems over an iSCSI-based storage area network (SAN), helping simplify both server management and the creation, distribution, and maintenance of server images. Now, Broadcom and Dell have teamed up to offer enhanced iSCSI boot functionality on Broadcom Ethernet controllers in Dell PowerEdge servers.

Networked storage in the data center

Using servers on a network to store data is not a new concept, nor is booting a server from a network. Network attached storage (NAS) has existed since the introduction of Network File System (NFS) more than two decades ago. NAS uses file-level protocols that abstract the underlying drive file system to provide networked storage such as shared drives in Linux® or UNIX® operating systems using NFS or in the Microsoft® Windows® OS using Common Internet File System (CIFS). Although NAS is well suited for file sharing, it does not provide block-based functionality, which includes the ability to boot an OS.

Technologies such as Preboot Execution Environment (PXE), Remote Program Load (RPL), and Bootstrap Protocol (BOOTP), which have existed for years, are the current mechanisms for booting directly from a network and are available on Broadcom Ethernet controllers and Dell PowerEdge servers. Although systems can be booted through a combination of NAS and PXE, RPL, and BOOTP, configuring systems to do so is not a simple matter, and may restrict system performance or capabilities.

The introduction of SANs provided block-level storage over a network, allowing shared resources to store application data from multiple servers and thereby helping increase storage utilization while reducing the number of devices for administrators to back up and manage. In this configuration, the file system appears to be present on a local direct attach storage (DAS) disk within the server, but actually exists on a SAN array. SAN technologies such as iSCSI help significantly simplify booting an OS from a

> "iSCSI offers a simple, cost-effective, standards-based way for even enterprises with modest budgets to deploy networked storage and gain the advantages of block-level storage consolidation."

network by providing the OS with low-level access to the storage device, which, in turn, allows native file system access and the use of standard OS disk utilities.

Networked storage systems such as NAS and SANs allow organizations to easily divide and share highly available pools of storage among groups of users, helping meet applications' increasing storage needs. These types of storage enable increased utilization, scalability, and simplified backup solutions while helping increase data availability. Multiple enterprise requirements contribute to the need for networked storage:

- **Efficient management:** Consolidating to networked storage allows administrators to manage a single pool of central storage resources, rather than disparate, unconnected storage systems. Consolidation can also provide benefits such as reduced total cost of ownership and easy retargeting of servers, helping create a dynamic, scalable data center.
- **High availability and robust disaster recovery:** Networked storage supports high levels of availability while helping simplify disaster recovery, backup, and data redundancy systems.
- **Regulatory compliance:** Government regulations such as the Sarbanes-Oxley Act and Health Insurance Portability and Accountability Act (HIPAA) have increased enterprise documentation, data security, and auditing requirements. Networked storage can help ease the task of complying with these regulations.

iSCSI SAN architecture

As shown in Figure 1, iSCSI-based SANs comprise two parts: an iSCSI initiator and an iSCSI target. The iSCSI initiator resides on the server or client system and connects to the iSCSI target over an IP network, allowing the physical disk on the target system to appear as a local disk. The iSCSI initiator sends commands to and receives responses from the iSCSI target; conversely, the iSCSI target receives and responds to commands from the iSCSI initiator.

The iSCSI protocol stack consists of a set of standard protocols residing at different Open System Interconnection (OSI) layers: Ethernet as the network interconnect (Layer 2), IP for routing (Layer 3), and TCP for transport (Layer 4), with the iSCSI protocol itself residing at Layer 5 (see Figure 2). iSCSI also allows a SCSI interface to be presented to the OS, simplifying OS and application utilization. In addition, because iSCSI is built on standard TCP/IP protocols, it can be routed over IP version 6 (IPv6) on packet-switched internetworks.

By combining existing and widely deployed standards, iSCSI offers many advantages over other SAN fabrics, including the following:

- **Reduced learning curve:** Standard TCP/IP and Ethernet components are typically familiar to IT staff, helping reduce training, administrative, and maintenance costs.
- **Interoperability:** Utilizing standardized networking technologies helps ensure interoperability between SAN components.
- **Flexibility:** iSCSI supports a wide range of Ethernet network topologies, which can take

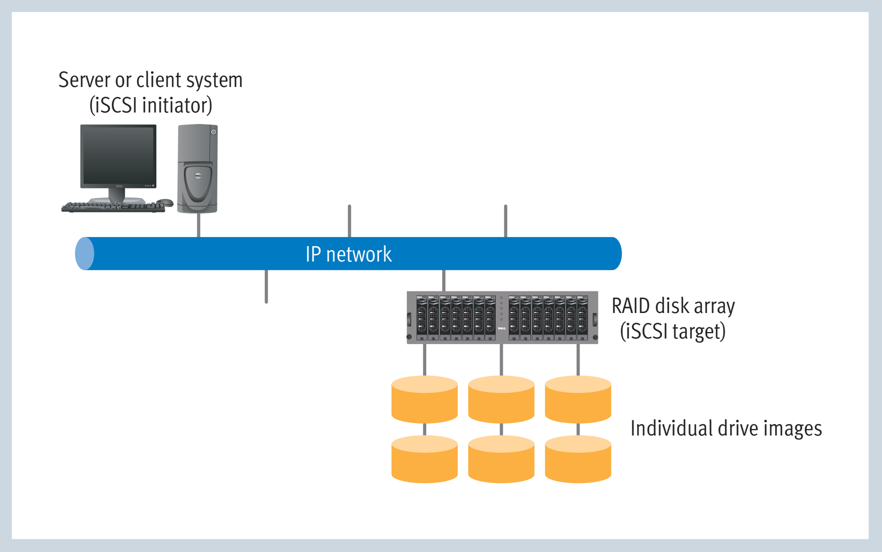

Figure 1. *iSCSI-based SAN architecture*

Protocol	OSI layer	Publication date	Specification
Ethernet	Layer 2	1974	IEEE 802.3
IP	Layer 3	1981	IETF RFC 791
TCP	Layer 4	1981	IETF RFC 793
iSCSI	Layer 5	2004	IETF RFC 3720

Figure 2. *Protocols used for iSCSI boot*

advantage of the routing capabilities inherent in the IP protocol. In addition, as a routable networked storage system, an iSCSI-based SAN can be used over the Internet and across the enterprise.

- **Cost-effectiveness and simplified management:** Using a single fabric for networking and storage helps avoid the additional acquisition and maintenance costs and administrative overhead of managing a separate fabric dedicated to storage.
- **Virtual LAN (VLAN) support:** VLANs help isolate network and storage traffic and support quality-of-service protocols on Ethernet.

These advantages have combined to accelerate the adoption of iSCSI—for example, iSCSI initiators are now a standard part of both Microsoft Windows and Linux operating systems.

iSCSI boot functionality over SANs

iSCSI boot allows a server to boot an OS over a SAN, helping eliminate the need for local disk storage, enhancing system reliability, and simplifying administrator workloads by centralizing the creation, distribution, and maintenance of server images. iSCSI boot can also accelerate server restore processes following a server failure: rather than having to re-image the replacement server, administrators can simply point the new server to the boot image on the SAN to quickly bring it online.

Deploying iSCSI to boot an OS over a SAN can be a relatively simple process. It requires the following system components:

- iSCSI boot–capable Ethernet controller with iSCSI boot ROM supporting the iSCSI, SCSI, IP, and TCP protocols
- iSCSI boot ROM to hook the BIOS Interrupt 13h interface and provide Extended Interrupt 13h interface services complying with Enhanced Disk Drive Specification 3.0, to allow support for drives larger than 2 GB through the use of 64-bit logical block addressing
- BIOS Boot Specification (BBS) support to allow the iSCSI boot implementation to seamlessly integrate with the BIOS boot process and boot selection interface
- OS-based iSCSI initiator, such as the Microsoft iSCSI Software Initiator for Microsoft Windows or the Open-iSCSI Software Initiator for Linux

The actual iSCSI boot process is divided into two phases: real mode and protected mode. These modes refer to the x86 processor operating mode, and roughly equate to the system BIOS and OS, respectively. In real mode, iSCSI boot transactions in systems with Broadcom Ethernet controllers are handled by the Broadcom iSCSI boot initiator. The BIOS uses this initiator to load the disk's master boot record (MBR), the OS bootstrap loader, and the OS bootstrap loader, then uses the Broadcom iSCSI boot initiator to load the OS kernel and the necessary OS drivers. Once the OS kernel is initialized and running, the processor transitions from real mode to protected mode and begins using the OS native drivers, including the OS iSCSI initiator, to complete the iSCSI boot process.

One exception to this boot process occurs when booting the Microsoft MS-DOS® OS using iSCSI boot. Because MS-DOS does not use the processor's protected mode, MS-DOS continues using the Broadcom iSCSI boot initiator for disk transactions even after it has booted.

One of the challenges of the iSCSI boot process relates to passing the iSCSI boot parameters and network settings from the Broadcom iSCSI boot initiator to the OS iSCSI initiator. Proprietary solutions have existed for some time, but today, the industry is starting to adopt the Microsoft iSCSI Boot Firmware Table (iBFT) specification, a table of iSCSI and network parameters written in memory by the iSCSI boot initiator and read by the OS iSCSI

"iSCSI boot allows a server to boot an OS over a SAN, helping eliminate the need for local disk storage, enhancing system reliability, and simplifying administrator workloads by centralizing the creation, distribution, and maintenance of server images."

"Broadcom and Dell have teamed up to create a simple yet richly featured iSCSI boot solution designed to allow iSCSI-based SANs to replace many forms of local storage."

initiator when it initializes. The table serves as a "handoff" between the real-mode and protected-mode phases of the boot process. Microsoft supports iBFT on the Microsoft Windows Server® 2003 OS and the next version of this OS, Windows Server 2008 (code-named "Longhorn"). Open-iSCSI supports iBFT in Linux.

Broadcom iSCSI boot solution on Dell PowerEdge servers

Broadcom and Dell have teamed up to include the Broadcom iSCSI boot solution on Dell PowerEdge servers that use the Broadcom NetXtreme II™ BCM5708 Ethernet controller. This feature is also available on Broadcom NetXtreme II BCM5708–based mezzanine network interface cards for Dell blade servers as well as on Broadcom NetXtreme™ BCM5722 Ethernet controllers.

Key features supported by Dell and Broadcom using iSCSI boot include the following:

- **Single ROM image:** The Broadcom iSCSI boot solution uses a single ROM image to provide iSCSI boot along with PXE, RPL, and BOOTP, helping avoid the need to flash separate ROM images when deploying different network boot technologies.
- **Integrated BBS support:** Booting from an iSCSI disk is as simple as selecting the Broadcom Ethernet controller from the boot menu in the Dell server BIOS.
- **Multiple target support:** Administrators can specify both a primary iSCSI target and a secondary failover target for iSCSI boot, so that if the server cannot establish a connection to the primary iSCSI target, it can use the secondary target instead.
- **Challenge Handshake Authentication Protocol (CHAP) support:** Support for one-way and mutual CHAP authentication helps increase security when establishing a connection between iSCSI initiators and targets.
- **Flexible iSCSI and network parameters:** When configuring iSCSI initiator and target information as well as standard network settings, administrators can either set the parameters statically in the Ethernet controller setup menu or retrieve them dynamically through Dynamic Host Configuration Protocol (DHCP).
- **iSCSI target redirection:** iSCSI targets can redirect the Broadcom iSCSI boot ROM to use a different iSCSI target for the boot process.
- **Linkup time delay:** Administrators can insert a delay during the boot process to accommodate targets that may need additional time to come online after a reset.
- **Microsoft Multipath I/O (MPIO) and device mapper (DM) multipath support:** Broadcom multipathing supports native MPIO in Microsoft Windows and DM multipath in Linux, helping provide network redundancy between iSCSI initiators and targets.
- **Target connections without boot:** Servers can make connections to iSCSI targets without actually booting from the iSCSI boot drive, helping make these targets readily available for direct OS installations.

Simple, flexible networked storage environment

As the use of SANs continues to grow, and the advantages of moving from local storage to centrally managed networked arrays become increasingly important, network boot solutions such as iSCSI boot should become a common feature of enterprise data centers. Broadcom and Dell have teamed up to create a simple yet richly featured iSCSI boot solution designed to allow iSCSI-based SANs to replace many forms of local storage. Deploying this solution can help enterprises of all sizes realize the advantages of iSCSI-based networked storage while helping simplify management and reduce total cost of ownership.

Robert Lusinsky *is a marketing director for the High Speed Controller Business Unit at Broadcom. Robert has a B.S. and an M.S. in Electrical Engineering from California State University and an M.B.A. from the University of Southern California.*

David Christensen *is an applications engineering manager for the High Speed Controller Business Unit at Broadcom. David has a B.S. in Computer Science from California State University, Fullerton, and has worked at a variety of software engineering and product support roles at AST Computer, Phoenix Technologies, and Broadcom over the last 17 years.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Broadcom:
www.broadcom.com

Dell PowerEdge servers:
DELL.COM/PowerEdge

Dell iSCSI storage solutions:
DELL.COM/iSCSI