



APPLICATION NOTE 4239

# Advanced Enterprise Features Enable Next-Generation SAS Systems

By: Sam Barnett, Business Manager

*Abstract: Designed for next-generation Serial-Attached SCSI (SAS) systems, Maxim's NexSAS™ product family is divided into four functional areas: expander technology, SAS/SATA support devices, server/enclosure/baseboard management, and signal conditioning. This article discusses the technology breakthroughs behind generation 1.5 NexSAS expanders.*

A similar version of this article was featured in [Maxim's Engineering Journal](#), vol. 63 (PDF, 1.3MB).

## Introduction

As first-generation systems based on Serial-Attached SCSI (SAS) technology enter the market, leading server and storage OEMs are racing to deliver the next group of platforms. In conjunction with the release of generation I systems, Maxim has enhanced its popular NexSAS™ product family.

Delivering high-performance storage solutions for SMB (small-/medium-business) and enterprise systems, generation 1.5 of the NexSAS family includes features found only in high-availability, superior-performing, Fibre Channel systems today. The NexSAS line comprises three new high-PHY-count expanders, extensions in intelligent mux/demux devices, a new enclosure-management/SAS-backplane controller, and the industry's best rate-agile signal conditioners.

Maxim's NexSAS product family is divided into four functional areas: expander technology, SAS/SATA support devices, server/enclosure/baseboard management, and signal conditioning. This article details the technology breakthroughs behind generation 1.5 NexSAS expanders. Specifically, it discusses the NexSAS multi-affiliation STP/SATA bridge and support for SAS 2.0 zoning.

## NexSAS Expander Technology

Designed for high-performance interconnect, the NexSAS expander family features a single functional design. This design approach ensures scalability from low-port-count offerings to the high-port-count devices needed for blade servers, storage enclosures, and switching/island SAN applications.

All of Maxim's expanders divide the core systems function into two separate components: the expander function and the management function. Included in the expander functional block are the expander connection manager (ECM), expander connection router (ECR), broadcast primitive processor (BPP), and the physical interfaces (PHYs) to the device (ranging in number from 6 to 36, depending on the application).

Complementing the core functions is the management functional block, which provides expander-management and vendor-specific enclosure-management functions. This block also integrates various peripheral interfaces needed for interaction in the overall system, including two-wire serial, JTAG, UART, general-purpose I/O (GPIO), and serial peripheral interfaces. The number and types of these interfaces vary with the particular expander and specific operating mode(s) employed.

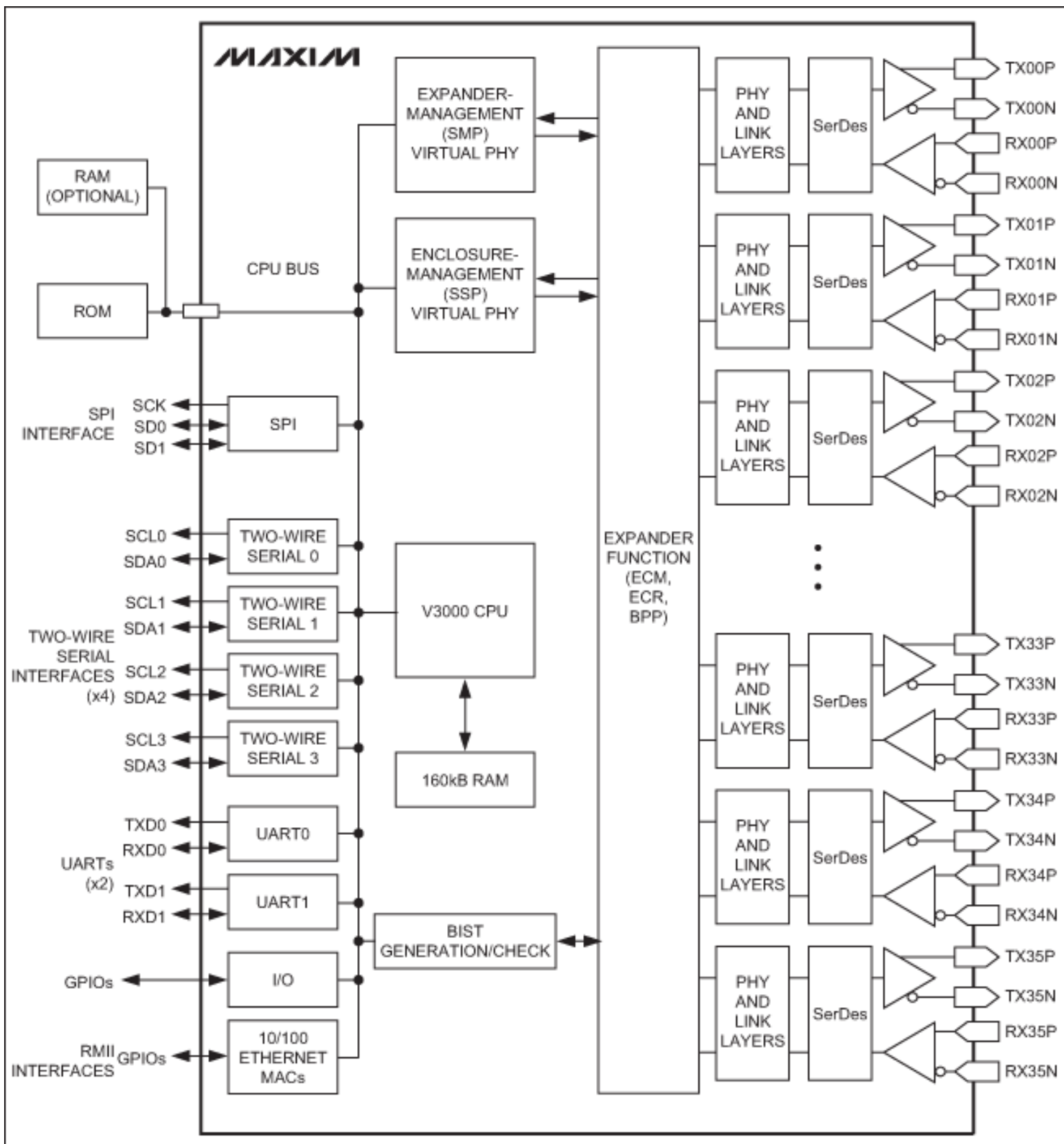


Figure 1. Functional diagram of Maxim's generation 1.5 NexSAS expander family.

**Figure 1** outlines the basic functional block diagram of all expanders in the NexSAS family. Generation 1.5 expanders offer a range of features to meet different system requirements. They also share a common set of features, including:

- **High-Performance Switching Architecture.** A low-latency, nonblocking switching matrix provides up to 54Gbps of aggregate switching capacity (in high-port-count expanders with linear scaling for lower-port-count expanders).
- **Maxim's Universal PHYs.** Self-configuring, rate-adaptable PHYs support SAS initiators and SAS/SATA targets at both 1.5Gbps and 3.0Gbps data rates. Multiple per-PHY preemphasis/deemphasis capability, and multiple per-PHY output levels leverage over 20 years of Maxim's technology development. They can be combined into SAS "wideports" of up to

*n* PHYs wide.

- **Virtual Initiator and Target Capability.** SMP (Serial Management Protocol), SSP (Serial SCSI Protocol), and STP (Serial ATA Tunneling Protocol) initiator and target functions enable robust value-added services, such as self-configuration, self-discovery, SES-mastering, and rogue drive identification.
- **Integrated Enclosure-Management (EM) Subsystem.** Robust enclosure-management capabilities maximize OEM investment in legacy firmware/software/custom features, and allow flexibility in system designs.
- **Integrated Ethernet Interfaces.** Two 10/100 Ethernet MACs are included in the EM subsystem to support blade servers or other applications that require Ethernet access to the expander.
- **Multi-Affiliation STP/SATA Bridge.** Extensions of the original STP/SATA bridge provide for two concurrent (active/active) NCQ affiliations between initiators and targets using Maxim's expander technology. This innovative approach eliminates the initiator starvation issues commonly associated with the single-affiliation mechanisms in SAS 1.1.
- **End-to-End SAS 2.0 Zoning.** For security, control, and performance, NexSAS expanders support both SAS 2.0 PHY-resolved and address-resolved zoning.
- **Other Enhancements Beyond Generation I Devices.** Additional capabilities include EPOW, integrated A/D converters, LPC for trusted-platform module interconnect, and parity protection of the on-chip memory.

Recent additions to the NexSAS family include the [VSC7156](#) 24-PHY, [VSC7157](#) 36-PHY, and [VSC7158](#) 18-PHY expanders. Like their predecessors (the [VSC7153](#), [VSC7154](#), and [VSC7155](#) generation I expanders), these devices are targeted at enterprise server, storage enclosure, and blade applications.

The above feature additions bridge the gap between the SAS 1.1 and planned SAS 2.0 specifications. Two of the key differentiating features, the multi-affiliation STP/SATA bridge and SAS 2.0 zoning, are examined in more detail below.

## Understanding the NexSAS Multi-Affiliation STP/SATA Bridge

### Fairness and Performance Limitations of Prior Designs

The SAS 1.1 specification defines the STP/SATA bridge as the expander component that provides a bridge between STP-capable initiators and SATA targets (HDDs). Without an STP/SATA bridge, SAS could not support Serial ATA (SATA) drives.

In general, the SAS 1.1 specification defines the function of the STP/SATA bridge as follows:

- In SAS expanders, the STP/SATA bridge allows the connection of SATA drives to the SAS domain.
- The SAS protocol is used for connection setup/tear-down between the host (initiator) and the STP/SATA bridge.
- The STP/SATA bridge only passes native SATA protocol to the SATA drive during connection.
- Expanders typically have one STP/SATA bridge available in each PHY, but the bridge is enabled only when connected to the SATA device.

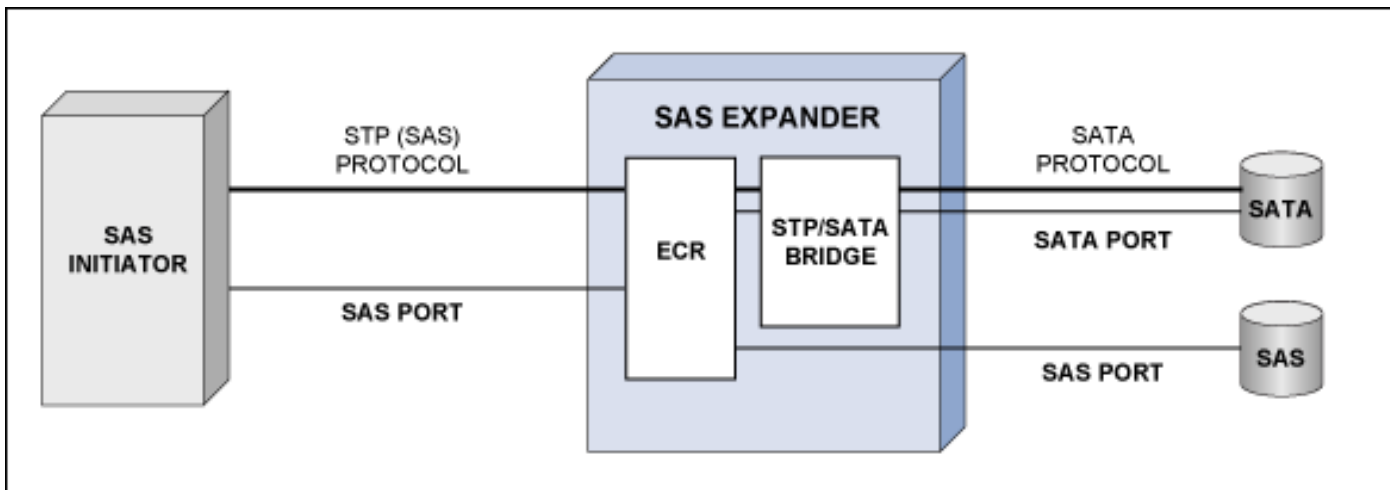


Figure 2. The STP/SATA bridge supports connections between STP-capable initiators and SATA drives.

**Figure 2** illustrates the STP/SATA bridge concept. Unlike SAS drives, SATA drives do not recognize multiple-host or multiple-initiator concurrent access. During a session, one STP initiator maintains an affiliation with the SATA target. During this affiliation, any other STP initiator attempting to gain access to the same SATA device receives an OPEN\_REJECT (STP RESOURCES BUSY) and is forced to retry the connection attempt. This simple affiliation mechanism maintains coherence by allowing commands from only one host to be active on the drive at any given time.

When all outstanding commands from a given initiator have completed, the initiator *should* send a special CLOSE (CLEAR AFFILIATION) primitive to the STP/SATA bridge, resulting in the release of the affiliation and allowing other hosts to access the

drive/resource.

As you can imagine, generation I expander devices support only a single affiliation in compliance with the SAS 1.1 specification.

With any new technology, implementation tends to expose certain undesired behaviors. The original STP/SATA bridging mechanism was no different. Limitations of the originally envisioned bridge involve both fairness and performance. Manifestations of these limitations are outlined briefly below.

### Fairness

- There are no formal limits imposed on the longevity of a given affiliation, meaning that an affiliation could be established forever.
- Fair access to SATA devices must be enforced by the initiators; in the absence of host-to-host communications, this may not be assured.
- Software algorithms to achieve fairness have proven painful for OEMs.
- Existing real-world STP initiators do not "play fair:" they occasionally fail to release affiliations.
- If a valid affiliation exists between the expander STP/SATA bridge and an initiator, other hosts cannot perform any inquiries to the drive, including diagnostic ones.

### Performance

- Single-threaded host access to SATA disks is slow in load-balancing systems.

Figure 3 depicts the SATA affiliation concept in SAS.

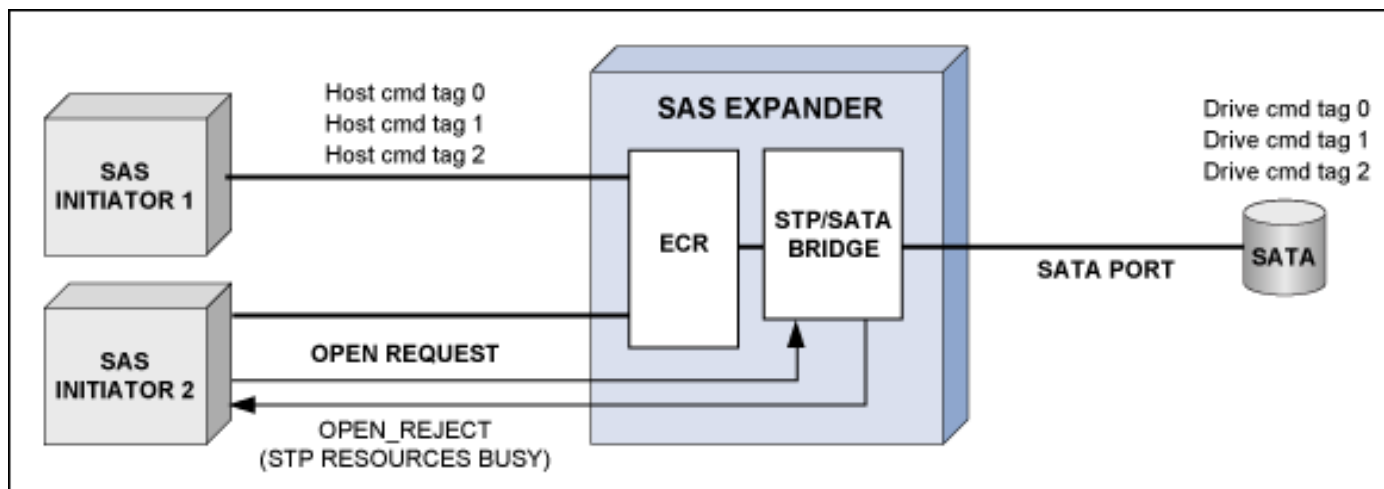


Figure 3. The SATA affiliation concept in SAS 1.1 only supports single affiliations, resulting in unanticipated fairness and performance limitations.

## New Approach Overcomes Traditional Performance Limitations

Generation 1.5 NexSAS expander products address both the fairness and performance limitations of the SAS 1.1 STP/SATA bridge. Through the updated bridge function, two hosts can simultaneously issue active commands on the same SATA drive.

Figure 4 illustrates the operation of the multi-affiliation STP/SATA bridge.

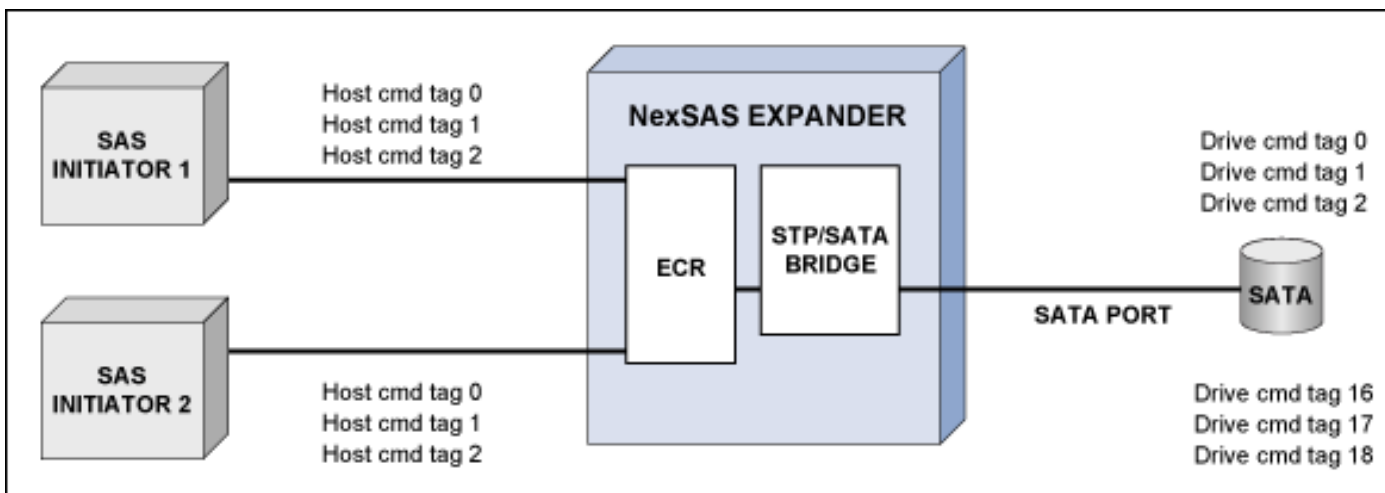


Figure 4. Generation 1.5 NexSAS expanders support multi-affiliation STP/SATA bridge operation, allowing two hosts to simultaneously issue active commands on the same SATA drive.

Details of the new bridge operation are summarized below.

### Intelligent Connection Management

- Flexible, nonexclusive disconnect policies (overlaps are allowed) include:
  - Timed disconnect (multiple modes and ranges)
  - Disconnect after any FIS transfer
  - Disconnect after data FIS transfer
  - Disconnect after interlocked operations
- Automatic callback of the proper initiator (host) based on queue tag/context

### NCQ and PIO Command Support

- Allows utilization of all 32 NCQ commands tags (or fewer by configuration) on the SATA drive
- Manages interleaving of NCQ and PIO commands from different hosts
- Tags are transparently managed and remapped to/from the drive to avoid tag number collisions between initiators (no special host software or firmware is required)

The multi-affiliation STP/SATA bridge employed by all generation 1.5 NexSAS expanders enhances fairness, performance, and system design convenience. Similarly, it benefits any SAS domain using a SATA disk that is shared by two or more hosts. It features an architecture that can scale for future products requiring even more active affiliations, while remaining completely compliant to the SAS 1.1 and proposed SAS 2.0 specifications.

## Understanding the Basics of SAS 2.0 Zoning

As storage OEMs and integrators embarked on the introduction and deployment of SAS-based systems, many longed for the days of Fibre Channel and some of the enterprise features it offers for traffic management (performance) and security.

To address those demands, the concept of end-to-end zoning was introduced into the proposed SAS 2.0 specification (refer to [www.t10.org](http://www.t10.org) for the latest on the zoning specification).

The SAS protocol, as defined, supports up to 16,384 devices (initiators, targets, expanders, and/or virtual devices) in a given domain; however, practical implementation of such large topologies is quite problematic. In essence, end-to-end zoning provides the mechanism necessary to manage these large physical topologies in smaller logical (segmented) groups. SAS zoning provides this function as well as the mechanisms to control access among and between groups.

The proposed SAS 2.0 specification divides zoning into two areas: PHY-resolved zoning and address-resolved zoning. Both approaches use identical access-control methodologies, but differ in zone group allocation.

### PHY-Resolved Zoning

In a PHY-resolved zoning implementation, each PHY of a zoning-capable expander is assigned to a zone group; any device attached to that PHY (or port, in the case of a PHY grouping for wideports) becomes part of that zone group. A permission table in the expander contains access-control permissions that outline the protocol for sessions between devices in differing zone

groups (by default, all devices within a zone group can interact with each other).

PHY-resolved zoning is ideal for smaller topologies, blade servers, or clustering applications.

## Address-Resolved Zoning

Address-resolved zoning provides a layer of abstraction not present in PHY-resolved zoning. In an address-resolved scheme, self-discovering expanders interpret each device address and populate a table. This table contains routing information and zone permissions. Permissions in the table and zone group assignment are combined to determine what level of interaction any two devices may have with each other.

Typically, address-resolved zoning is used for larger topologies or where more granular control is needed.

## Conclusion

One size never fits all, be it SAS or something else. Proper solutions differ by application, connectivity requirements, scalability, performance, and price sensitivity—all functions of innovation. With this in mind, Maxim's wide portfolio of storage networking products offers designers unmatched flexibility in choosing the feature sets that are ideal for their design challenges.

For a list of all the products in the NexSAS storage products family—the server-series expanders, the enclosure-series expanders, the SAS/SATA support devices, and the backplane/enclosure-management/baseboard-management controllers—go to [Storage Products](#).

A similar article appeared in the January 2008 edition of the *Serial Storage Wire* online newsletter published by the SCSI Trade Association.

NexSAS is a trademark of Maxim Integrated Products, Inc.

---

Application Note 4239: [www.maxim-ic.com/an4239](http://www.maxim-ic.com/an4239)

### More Information

For technical support: [www.maxim-ic.com/support](http://www.maxim-ic.com/support)

For samples: [www.maxim-ic.com/samples](http://www.maxim-ic.com/samples)

Other questions and comments: [www.maxim-ic.com/contact](http://www.maxim-ic.com/contact)

---

### Automatic Updates

Would you like to be automatically notified when new application notes are published in your areas of interest? [Sign up for EE-Mail](#).

---

### Related Parts

VSC7154: [QuickView](#)

VSC7155: [QuickView](#)

VSC7156: [QuickView](#)

VSC7157: [QuickView](#)

VSC7158: [QuickView](#)

VSC7160: [QuickView](#)

VSC7161: [QuickView](#)

AN4239, AN 4239, APP4239, Appnote4239, Appnote 4239

Copyright © by Maxim Integrated Products

Additional legal notices: [www.maxim-ic.com/legal](http://www.maxim-ic.com/legal)