

Serial Storage Architecture



By Dave E. Hall

This article describes Serial Storage Architecture (SSA), a definition and general specification of a high-performance serial link for the attachment of input/output devices. An overview of the architecture is followed by a description of the IBM SSA products available for the RS/6000™ under AIX®, and some hints and tips on configuring SSA products with AIX.

Serial Storage Architecture (SSA) defines a high-performance serial link for the attachment of input/output devices. It has been optimized for storage applications such as hard disk drives, host adapter cards, and array controllers.

SSA has many advantages over existing parallel interfaces such as the Small Computer Systems Interface (SCSI-2). It uses compact cables and connectors, and it has better performance, connectivity, and reliability. SSA is an open standard ANSI® architecture. IBM first announced and started shipping products based on SSA for use with the RS/6000 in August 1995. It has subsequently announced support for SSA products on IBM Personal Computers and other manufacturers' systems, such as Hewlett-Packard® and Sun®.

By November 1997, IBM stated that it had shipped over two petabytes of SSA storage since its introduction, and that it was currently working on a future implementation of SSA-based products that would be twice the speed of existing ones. SSA products form a key part of the recently announced Seascape Architecture.

SSA Overview

The heart of any IBM SSA-based product is the SSA chip—the Serial Interface Chip (SIC). The SIC, implemented in IBM CMOS4-LP technology, provides two 20 MB/sec serial ports. A single SIC can be used to implement a dual-port SSA disk drive, host adapter, or bridge controller. For the products described in this article, each disk drive has one SIC and each adapter card has two.

The SSA design allows three basic configuration possibilities: string, loop, or switch. The most common one implemented in products is loop, because it has both higher bandwidth and reliability. Figures 1 and 2 show typical SCSI and SSA subsystems. Figure 3 shows a comparison of SSA and SCSI.

RS/6000 SSA Products Overview

Several products provide low-cost, high-performance disk storage for the RS/6000.

7131/7133 SSA Disk Subsystems

The 7131 is the entry-level SSA subsystem for the RS/6000 and RS/6000 SP™. It is sold as standard with a pair of high-performance Ultrastar disk drives; each pair consists of disks of 2.2 GB, 4.5 GB, or 9.1 GB capacity. Three additional hot swappable slots can contain a mix of any of these capacity drives for a maximum subsystem configuration of 45.5 GB. For storage requirements greater than this, multiple 7131s can be connected together.



Dave E. Hall

Typical SSA Subsystem

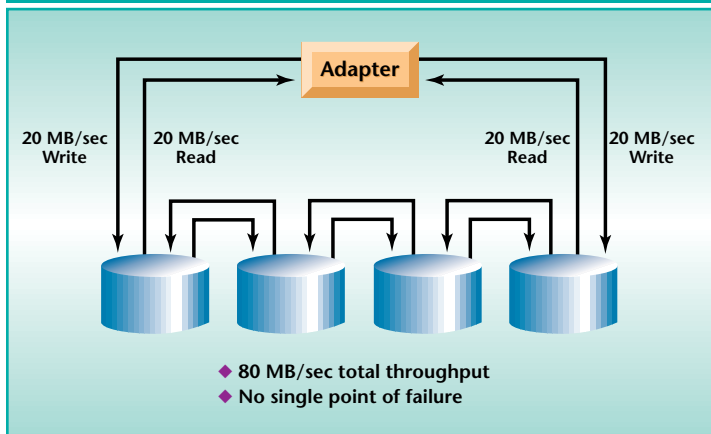


Figure 1. Typical SSA subsystem

Typical SCSI Subsystem

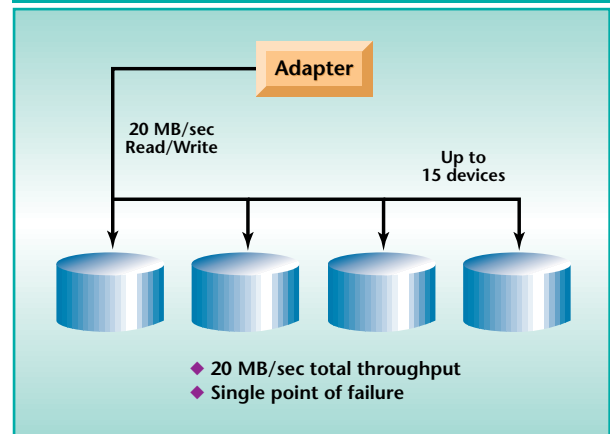


Figure 2. Typical SCSI subsystem

The 7133 is the highest performing disk subsystem for the RS/6000 and the RS/6000 SP. It is available as a rack mount or a desk-side tower unit. Its configuration is similar to the 7131, with Ultrastar drives that can be intermixed for flexibility in building storage environments from gigabytes to terabytes. Its serial loop architecture provides redundant paths to all disks. The 7133 supports mission-critical applications via High Availability Cluster Multiprocessing (HACMP).

Two adapters attach the 7131 and 7133 to multiple RS/6000s. The 4-Port Adapter connects up to two hosts while the Enhanced 4-Port Adapter can attach up to eight hosts. This allows multiple hosts to access the same data with high performance and availability features of SSA. Since SSA devices are configured in loops, they do not require bus arbitration. This enables multiple concurrent operations to

occur in separate sections of the loop, resulting in higher overall throughput.

Two adapters provide RAID-5 capability for the SSA serial interface. The 4-Port RAID adapter supports Micro Channel® systems while the PCI SSA 4-Port RAID adapter supports selected PCI systems. Both adapters are ideal for video applications, data servers, and mission-critical storage needs.

IBM 7190 SCSI Host to SSA Loop Attachment

As the number of storage devices increases on a SCSI bus, performance decreases dramatically. The 7190 overcomes these SCSI limitations by allowing users to efficiently attach high-performance 7131 and 7133 SSA disk storage subsystems to systems that do not provide native SSA attachment capabilities. Up to 48 SSA disks can be

SSA	SCSI
4 signal wires—6 total	31 signal wires—68 total
Up to 80 MB/sec system throughput	Up to 20 MB/sec system throughput
No discrete terminators	Discrete terminators
127 devices per loop or string	15 devices per string
Hot plugging (simple, designed in)	Hot plugging (possible but difficult)
Alternate paths avoid a single point of failure	SCSI bus is a single point of failure

Figure 3. Comparison of SSA and SCSI

connected to each SCSI adapter. Because each 7190 and its complement of SSA disks looks to the system like a single SCSI device, the level of arbitration is dramatically reduced and typically leads to significant increases in performance over the traditional SCSI attachment method. The 7190 is supported on IBM and selected Sun and Hewlett-Packard computers.

The 7190 uses a Differential Fast/Wide SCSI-2 bus adapter to connect to the system. Because it relies on the host's standard SCSI driver and hardware for communications, there is no need for modification to system hardware, software, or applications.

All attached storage subsystems can be monitored online for the following:

- ◆ Display vital product data
- ◆ Show SSA loop topology
- ◆ Show SSA-to-SCSI ID/LUN mapping
- ◆ Set or change modes for drive identification and diagnostics

- ◆ Download microcode to the 7190 and SSA disk drives as required
- ◆ Provide detailed system-wide error log
- ◆ Alert system to redundant power/cooling errors, loop topology changes, or drive errors

SSA and AIX Devices

SSA subsystems and adapters introduce some new terms for AIX. Figure 4 shows the result of an AIX `lsdev` command issued against some SSA devices.

Note from the example in Figure 4 that `pdisk0` is not necessarily associated with `hdisk0`; it is, in fact, associated with `hdisk1`. System administrators must determine the `pdisk`-to-`hdisk` mapping in order to perform certain functions, such as assigning disks to volume groups, cabling for performance, changing disk placement for performance, handling disk failures, and so on. Fortunately, with AIX 4.1.5 this

Command	State	Device
ssar	Defined	SSA Adapter Router
ssa0	Available 00-05	SSA Adapter
ssa1	Available 00-06	SSA Enhanced Adapter
ssa2	Available 00-07	SSA RAID Adapter
pdisk0	Available 00-05-P	1 GB SSA C Physical Disk Drive
pdisk1	Available 00-05-P	2 GB SSA C Physical Disk Drive
pdisk2	Available 00-06-P	2 GB SSA C Physical Disk Drive
pdisk3	Available 00-06-P	4 GB SSA C Physical Disk Drive
hdisk1	Available 00-05-L	SSA Logical Disk Drive
hdisk2	Available 00-05-L	SSA Logical Disk Drive
hdisk3	Available 00-06-L	SSA Logical Disk Drive
hdisk4	Available 00-06-L	SSA Logical Disk Drive

ssar—SSA router, which should always be in the defined state
 pdisk—physical disk
 hdisk—hard disk; one `pdisk` per `hdisk` for non-RAID, multiple `pdisks` per `hdisk` for RAID

Devices are configured in "walk the loop" order.

- ◆ Adapters from lowest numbered slot
- ◆ `Pdisk` in serial number order for each adapter
- ◆ `Hdisks` following the `pdisks`

Figure 4. Result of `lsdev` command on various SSA devices

SSA Domains

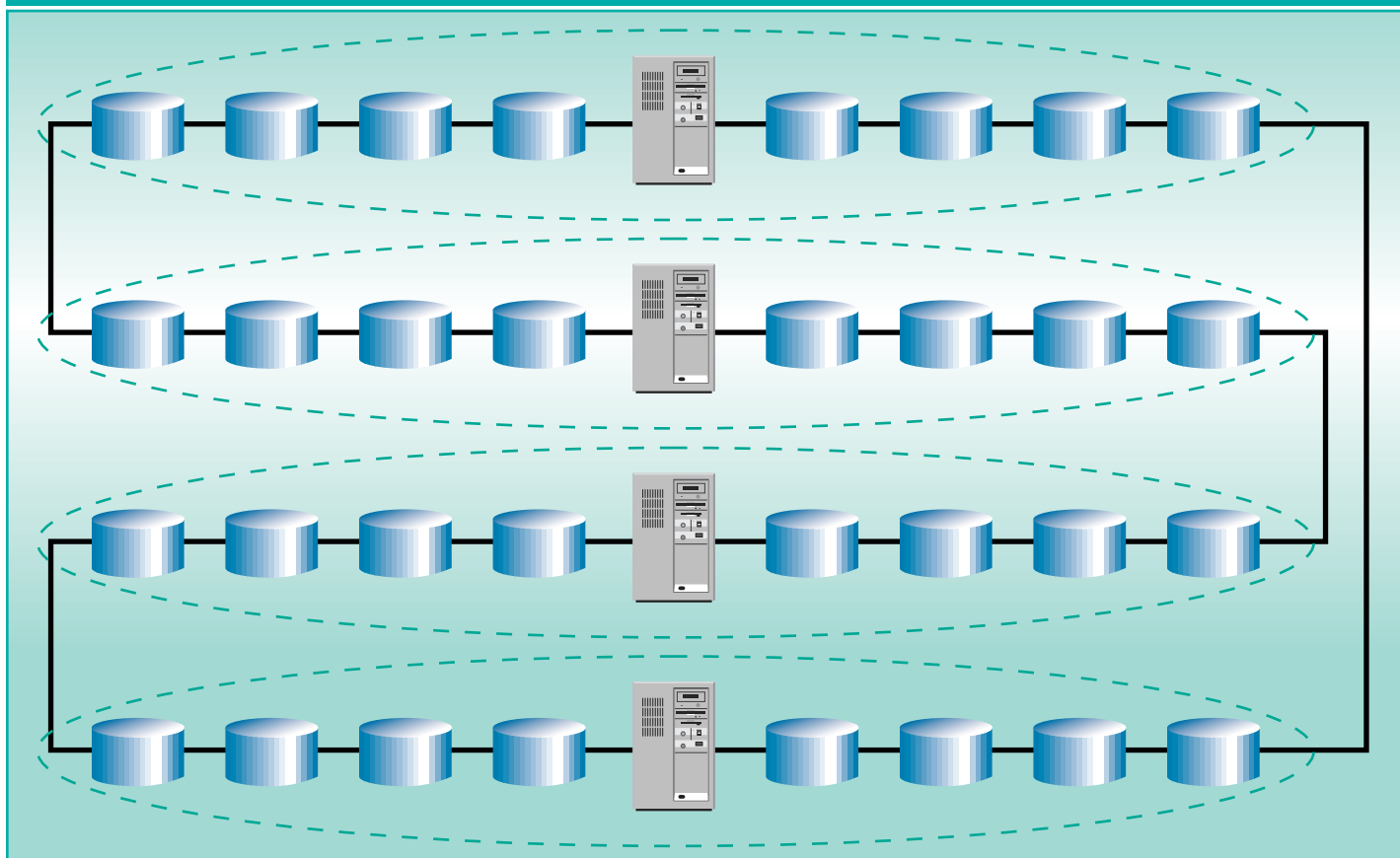


Figure 5. SSA domains

can be done via the System Management Interface Tool (SMIT).

The numbers associated with the adapters and disks are determined in the following way. When the system performs the “walk the loop” function at initialization, it configures the first SSA adapter that appears as `ssa0`, the next as `ssa1`, and so on. For each adapter, the SSA disks connected to it are configured in serial number order, as displayed on the front of each drive. This is not necessarily the order of the disks going around the loop.

SSA Domains

Figure 5 illustrates the concept of SSA domains; each domain is circled with a dotted line. Note that the adapter closest to a disk is in the same domain. The domain is defined from the adapter’s point of view. If

there are multiple adapters in a loop, it is best to put the disks that will be accessed by a particular adapter closest to it. Then, I/O from the disks to their adapter will not interfere with the I/O from other disks to their adapter.

The adapter port domain means that for a particular loop of disks, those nearest to a port will be in that port’s domain. For multiple computers in a loop, be sure that when you set up volume groups, the disks that comprise the group are closest to the system on which they are most likely to be used. This eliminates I/O interference between computers.

The loop arrangement enhances data availability by allowing one computer to access another computer’s disks without any recabling. For example, a volume group can be varied off-line from one computer and

varied online to another. This enhances data availability if a computer fails or must be taken off-line for maintenance.

Dave E. Hall, *IBM Corporation, Langstone Road, Havant, Hampshire PO9 1SA, England. Mr. Hall works in the SSA Architecture and Development organization. He has a BSc in Mathematics from Brunel University.*