Industry Association

Volume 1
June 26, 1996

SSA: Understanding the Facts

Why Customers are Buying SSA Solutions

As the debate over serial interfaces continues, one thing is perfectly clear: Customers are flocking to buy SSA storage solutions to meet their growing needs for:

1. High Performance
2. High Reliability & Availability
3. Low Cost
4. Products Available Today

This explosion in demand for SSA solutions is occurring today in several markets: Servers, High End Storage, Audio/Video Industry, Graphics Arts Prepress and High Performance Data Mining. The technology behind SSA's attributes will be examined in this article, as well as contrasting SSA to other competing technologies such as FC-AL Disk Profile.

In the Unix market, IBM has been shipping the 7133, the first serial disk subsystem, with general availability since August, 1995. In the first 8 months IBM shipped over 200 Terabytes of SSA storage to over 1,000 different customers. And, by the end of June 1996, over 300 Terabytes of SSA storage will be installed at a variety of end customers. SSA has shown the fastest ramp of any product in the Unix marketspace.

A number of other manufacturers are enabling SSA storage to be attached to non-IBM platforms this year, including Sun, HP, Apple and Unisys. Pathlight and Vicom are two such manufacturers. Vicom offers a bridge product which converts SSA to SCSI and vice versa. This allows SSA subsystems to be attached to existing SCSI channels, providing much higher performance to the end user than SCSI solutions offered today. Pathlight has brought SSA to the Audio/Video Market with their PCI to SSA adapters running on a number of platforms and operating systems, including MAC O/S and Windows NT. Pathlight also offers attachment to a variety of non-IBM servers by providing operating systems support for Novell, Windows 95, SCO Unix, and NeXTStep. A number of additional SSA component suppliers have been key to bringing SSA solutions to customers. A list of suppliers is available from the SSA IA.

At the National Association of Broadcasters conference (NAB) in April this year, there were 14 companies endorsing or displaying functioning SSA products including: Scitex Digital Video, Data Translation Inc., Discreet Logic, ASC Audio/Video, and Mitsumi. Enclosures were offered by Microtech International, IBM, Siemens/Nixdorf, Xyratex, Sigma Trimm Technologies, and Zentra. Why has the Audio/Video industry embraced SSA? SSA solutions offer the best performance for A/V applications at a low cost, and they're available today!

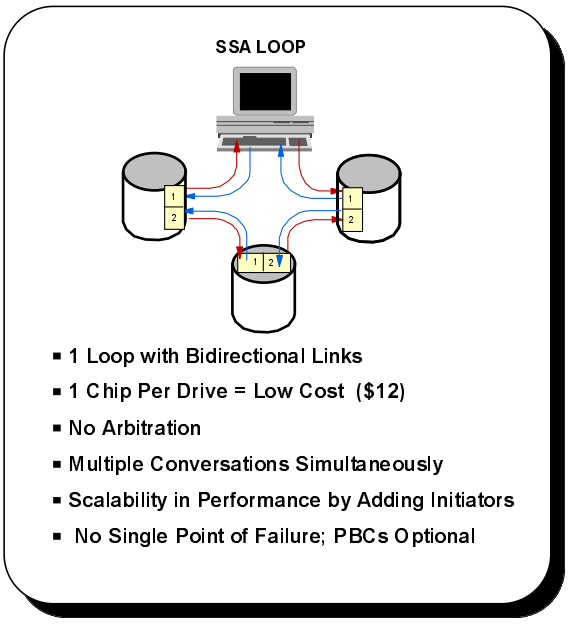
Advantages of SSA

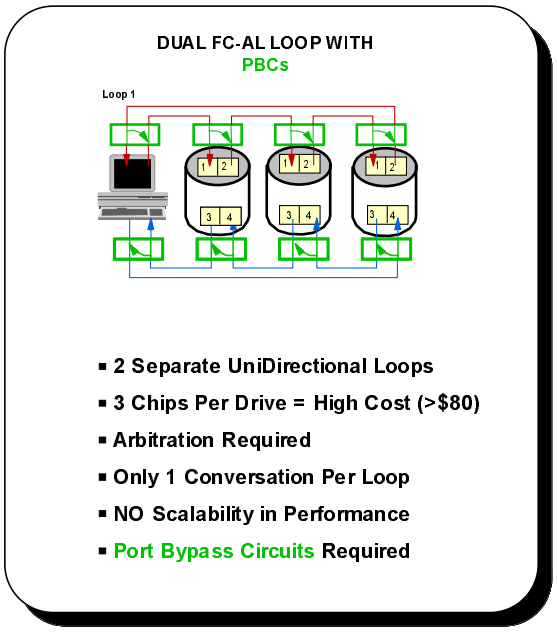
SSA has some unique advantages which make it the appropriate interface choice for attaching storage and peripherals to systems. The key advantages outlined in the following sections include:

SSA Advantages	***FC-AL Disadvantages***
1. Inherent Reliability & Availability Characteristics (No Single Point of Failure) due to SSA's full duplex, bidirectional links and dual ports.	***1. FC-AL requires active components on the backplane, Port Bypass Circuits, to implement no single point of failure.***
2. SSA allows for Scalable Storage Solutions which provide Higher Performance as more systems are added to an SSA loop.	***2. FC-AL Does Not allow for Scalability due to its arbitration characteristics and lack of spatial reuse.***
3. SSA provides Superior Fault Isolation and Error Recovery that is implemented on a frame by frame basis.	***3. FC-AL is still wrestling with how best to perform fault isolation and error recovery.***
4. SSA's ability to have multiple conversations on a loop ("Spatial Reuse") provides superior implementations for functions such as "XOR on the Drive" and "Disk to Tape Backup".	***4. FC-AL does not allow for simultaneous disk to tape back up on a loop due to its lack of spatial reuse. FC-AL's implementation of XOR on the drive is inferior to SSA's for the same reason.***
5. SSA is a low cost interface due to the ability to implement SSA as a single chip solution. The total cost is $40 for a single chip SSA initiator.	***5. FC-AL requires 3-4 chips to implement just the interface. This totals to a cost of $186 for an FC-AL initiator interface.***

Architecture

SSA and FC-AL are both loop architectures with packetized frames. However, the transmission and delivery of these frames and the basic architectures are considerably different. SSA was designed specifically for storage systems, with the necessary performance, reliability, error recovery and cost features that storage requires. FC-AL, and more specifically FC-AL Disk Profile, came out of a huge Fibre Channel Standard that tries to provide all services to all possible applications. FC-AL was originally designed as a networking interface, but is now being force-fit into a drive interface via the "FC-AL Disk Profile". Several problems result from this approach in the areas of performance, reliability and error recovery - which will be discussed in this article.





The figures above describe the basics of both the SSA and the FC-AL architectures. SSA is a non-arbitrated, full duplex, bidirectional loop. The basic building block in SSA is the "PORT". SSA ports are connected in a point to point manner by a full duplex, bidirectional "SERIAL CONNECTION". An SSA serial connection consists of two "LINKS", each of which is a 2-wire twisted pair. One link provides inbound data, the second link provides outbound data, each operating simultaneously and independently. SSA chips available today automatically come with 2 ports, with an internal router that allows data to pass through (i.e. between ports) with very low latency. The total bandwidth at an SSA node is 80 MB/S, 40 MB/s of read AND 40 MB/s of write. Since each link operates independently, read and write operations can occur at the same time on a loop. This is called "SPATIAL REUSE". Furthermore, since each serial connection provides a bi-directional path, there is no single point of failure in an SSA loop. If a node or connection should fail, the initiator will automatically reroute frames around the other side of the loop.

By contrast, FC-AL (Fibre Channel ARBITRATED Loop) is an ARBITRATED, half-duplex, unidirectional loop. An FC-AL serial connection is a single link providing transmission in one direction. An FC-AL port receives inbound data and forwards it on in the same direction. Data transmission down the serial link is 100 MB/s and therefore the maximum transmission at any device on the loop is 100 MB/s. There is no "Spatial Reuse" in FC-AL, thereby limiting the scalability and performance of FC-AL loops. Since frames in FC-AL are only sent in one direction around a complete loop, the loop must always remain unbroken for any read or write operation to occur. This means that if any of the connections or devices fails, the loop is unusable. Hence, the concept of "Port Bypass Circuits" or PBCs was created. PBCs are active components that plug into the backplane of an enclosure to allow a serial connection to "bypass" a failed device. This was a great invention for FC-AL, even with the added cost, because it now meant that a drive could fail and the loop would remain intact. The only problem was that having drives with single ports, single loops and single PBCs left SINGLE POINTS OF FAILURE. Enter "Dual FC-AL Loops". Dual loops basically doubles the Fibre Channel

Circuitry on a device and provides 2 counter rotating loops with PBCs for both ports of each device.

Reliability and Availability

As discussed in the architecture section of this article, SSA has inherent capabilities to provide no single point of failure without any additional circuitry due to its full duplex, bidirectional links. This feature provides for a number of characteristics in storage subsystems that are crucial, namely:

- Good Fault Isolation and Error Recovery
- Flexibility in configurations
- Hot Pluggable Drives
- Serviceability

SSA was designed specifically with storage solutions in mind. Therefore much attention has been paid to the way in which SSA handles fault isolation and performs error recovery. Unlike FC-AL Disk Profile, SSA detects faults (failures) on a link, connector, disk drive, or enclosure. This capability provides for excellent fault isolation characteristics. In FC-AL Disk Profile, failures cannot be isolated to a link or a drive. In FC-AL Disk Profile, when an error occurs the entire data transfer, including all frames, must be resent. One of the reasons that SSA is so well suited for storage, is its Error Recovery capabilities. In SSA, when a fault is detected, Error Recovery occurs at the lowest levels of the protocol, on a frame-by-frame basis. The value in this is that the SSA hardware and microcode handle error recovery procedures automatically between neighboring ports, isolating any problem down to a specific frame.

Having the capability to perform point to point error recovery in SSA means that users have greater flexibility in changing their configurations without impacting performance. To implementers, it means a clean and proven method of handling and recovering from errors. In FC-AL Disk Profile, the methods of performing error recovery have not even been fully determined yet.

Flexibility of configurations is key to providing customers with functions such as hot pluggable drives and easy serviceability. When an SSA drive is pulled from a loop, an ASYNC ALERT message is sent back to the initiator to indicate that there is a change in the configuration. The initiator will only rewalk the portion of the SSA loop which has changed to determine the changes to the configuration. The time for this is negligible, and the customer can still have access to the remaining drives. When the drive is plugged back in, automatic reconfiguration occurs again and the customer has access to the complete loop of drives.

For those customers that require protection against multiple drive failures, SSA allows the use of bypass circuits in a backplane environment. This would allow customers to remove multiple, non-neighboring drives from an enclosure without loss of access to any of the remaining drives. A number of manufacturers provide bypass circuits for SSA customers requiring this added protection for multiple points of failure.

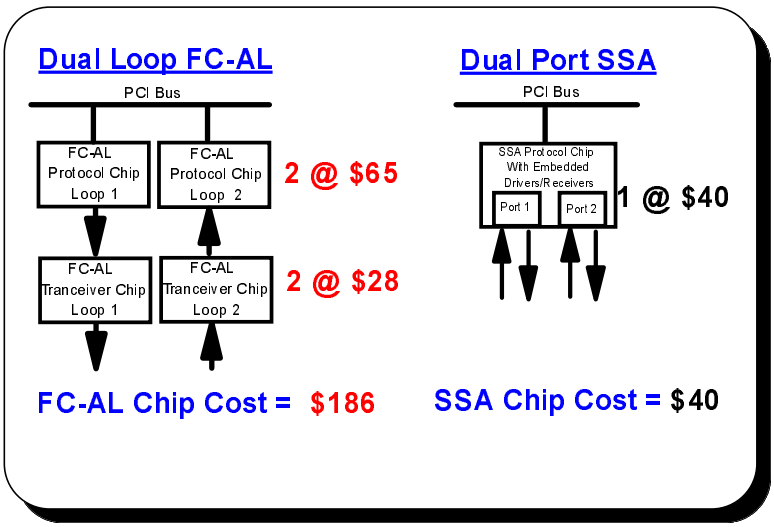
SSA: The Low Cost Serial Interface

One of the key benefits of SSA is its low cost implementation. An SSA target or initiator interface is implemented in a single CMOS chip, with 2 ports, each containing:

- Imbedded High Speed Drivers and Receivers
- 8B/10B Encoder/Decoder
- Serializer/Deserializer
- Phase Lock Loop
- Protocol State Machine
- Frame Buffers

Also included in the same chip is the Cut Through Routing which provides the capability of frames to pass through nodes with very low latency. This entire implementation, including all analog components comes to approximately 27K gates. It is implemented today in 3.3 Volt, 0.5 micron CMOS technology. The exact same circuitry can be used for both Target and Initiator SSA nodes. In fact, the SSA core is small enough to implement a single chip target (SSA core + Buffer Formatter) or a single chip initiator (SSA Core + Processor + Host Bus Interface). The single chip initiator exists today in an off- the-shelf component from VLSI Technology Inc., containing: Dual Port SSA Core, 256 DMA Channels, ARM Processor and PCI Interface!

FC-AL however requires 3 OR 4 CHIPS to implement JUST THE INTERFACE. This does not take into account any integration for microprocessors or disk controllers. The reason is that the high speed drivers and receivers require external components in FC-AL. So for a Dual Loop FC-AL interface one would require 2 driver/receiver chips (one per loop) and at least one, and more likely two, protocol chips. If you can find an FC-AL disk drive or adapter, ask about the number of chips used to implement both the FC-AL Physical (driver/receiver) layer and the protocol layer. For example, HP has announced their FC-AL protocol chip, as well as their Driver/Receiver chips. For a dual loop implementation, two of each of these chips are required to handle both loops. Below is a cost comparison between SSA and FC-AL to implement the interface for an initiator in each technology:



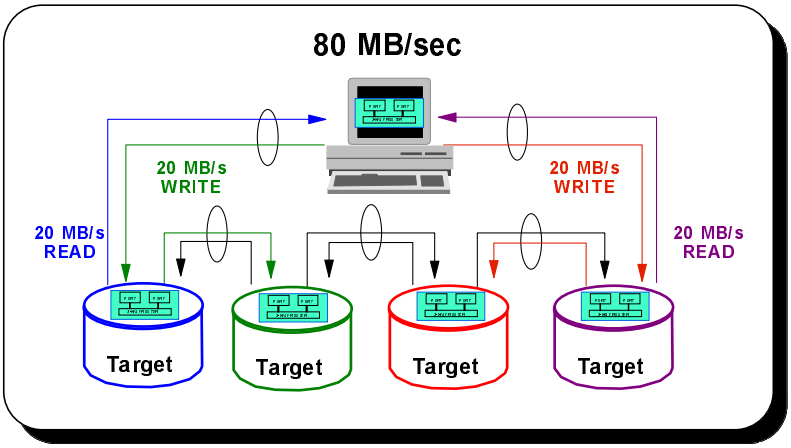
The Power of Spatial Reuse

Spatial Reuse, the ability to perform multiple conversations on an SSA loop, provides several benefits to storage subsystems, including:

1. The ability to scale performance by adding additional initiators to an SSA loop.
2. A superior solution for XOR on the drive due to SSA's capability of performing the XOR function without any involvement from the initiator, and while other operations are being performed on the loop.
3. A superior solution for Tape Back Up, again because SSA allows a tape backup or restore operation to occur without involvement from the initiator, and while the initiator is performing other operations on the loop.

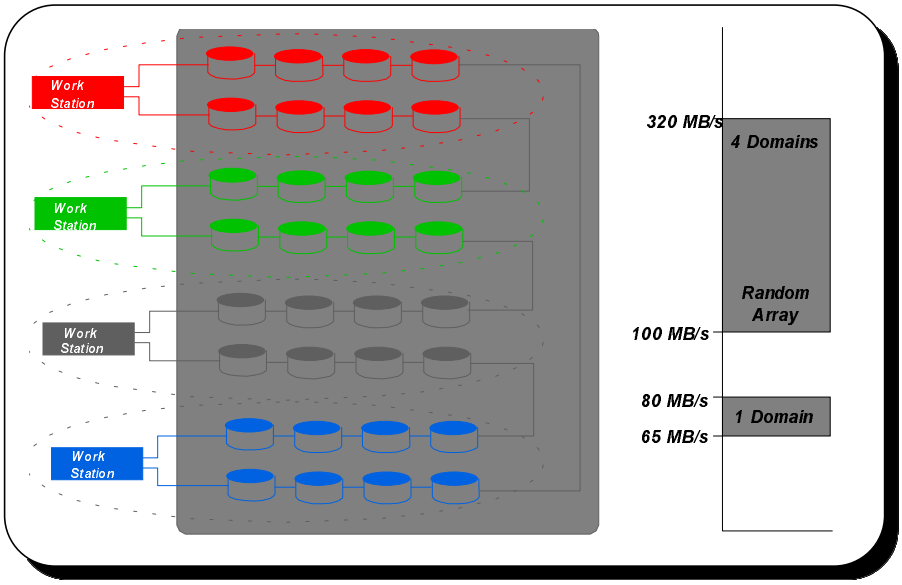
SSA Performance Scalability

An SSA loop with a single initiator delivers 80 MB/s of aggregate bandwidth by performing operations on all links simultaneously. The diagram below depicts how this is achieved.

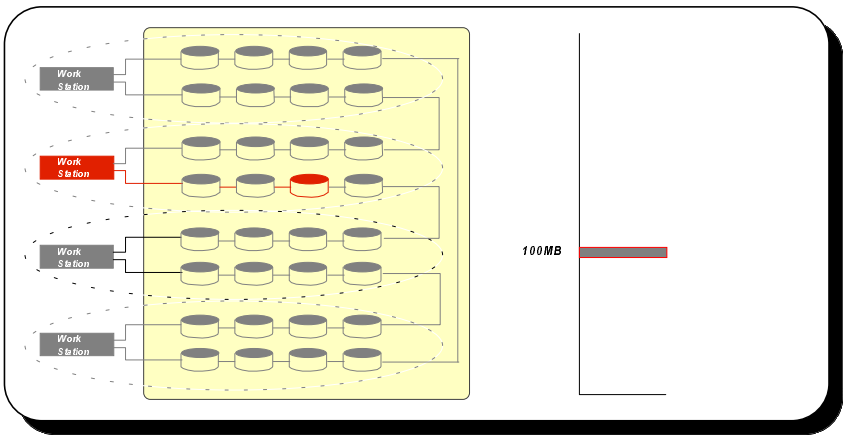


Performance scalability occurs in SSA by adding additional initiators to a loop. Each initiator has access to the drives within its own *domain* (i.e. those drives off either port of the initiator), as well as those in the domain of other initiators on the loop. Each initiator on the loop, therefore, adds an additional 80 MB/s of bandwidth to the loop when accessing data completely within its own domain. In the figure below, for example, with 4 initiators (servers) in an SSA loop, the total aggregate bandwidth can be as high as 4 x 80 MB/s, or 320 MB/s when all initiators are reading and writing within their own domain. This scalability is due to the effects of spatial reuse, the capability to perform multiple operations on an SSA loop simultaneously. In a real world environment where you may have a combination of drive access within a domain and completely random access to any of the shared drives on the loop, the aggregate bandwidth has been measured in

systems with 4 initiators at 109MB/s. *SSA storage subsystems are the only products where the total system bandwidth can exceed the bandwidth of the interface.*



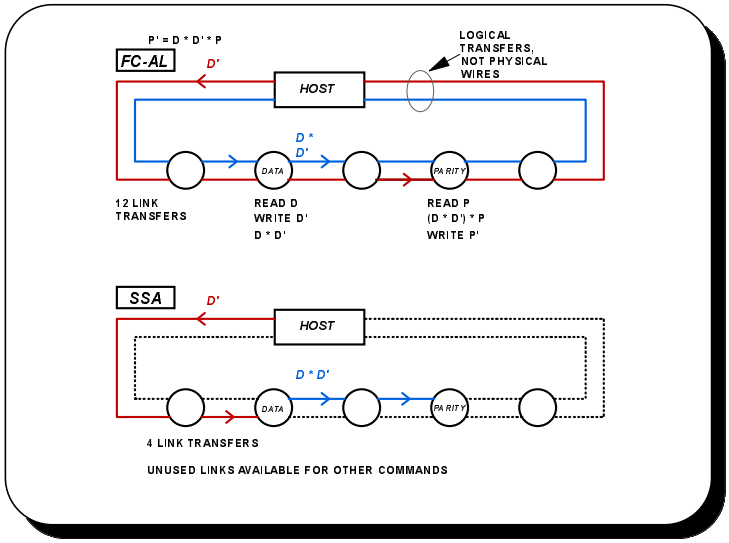
By contrast, FC-AL cannot scale performance due to the effects of arbitration. Only a single drive can be communicating with a single initiator on an FC-AL loop at one time. So, showing the same configuration in FC-AL would give you a maximum theoretical performance of 100 MB/s. This is depicted in the figure below.



XOR on the Drive with SSA

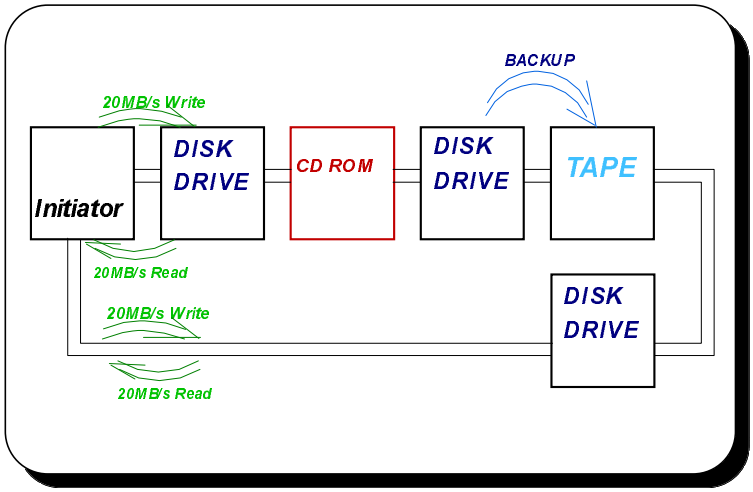
Another benefit of spatial reuse is the perfect fit for XOR on the drive capability. By taking advantage of the independence of the SSA links, an XOR configuration can be maximized by enabling the drives to perform the XOR operation without involvement from the initiator. This frees the initiator to perform operations on other drives in the loop, while the XOR function is being executed. The following diagram depicts the advantages of SSA is this enviroment over FC-AL, reducing the number of operations that must be performed by the initiator. In SSA only 4 link transfers occur for an XOR

function, but in FC-AL (again due to Arbitration) up to 12 link transfers must occur to execute the same function.



SSA for Tape Backup

A third area that maximizes the effects of spatial reuse is tape back up. Similar to the XOR on the drive function, tape back up implemented with SSA's spatial reuse techniques allows a disk drive to be backing up to tape while other operations are taking place on the SSA loop. This is a terrific advantage for cusotmers who need access to their data 24 hrs a day, and don't want to suffer down time while taking a system offline for back up. Again, this cannot be done in the same way with FC-AL because they can only perform a single operation on a loop. In FC-AL a tape back up operation would take place without the customer having access to other peripherals on the loop. The diagram below depicts tape backup occuring in SSA, while the initiator is accessing disk drives and a CDROM drive on the same loop.



Roadmap to Higher Performance

The SSA technology is not stopping with 20 MB/s links. The next version of the technology, 40 MB/s links, is already implemented in silicon. This will provide future SSA products with total bandwidth of 160 MB/s per node, doubling all performance numbers seen today. This faster version was achieved by moving to a 0.35 micron CMOS technology, with faster circuitry. The protocol essentially remains the same. Furthermore, once 40 MB/s link products are available, customer will be able to mix 20 MB/s link products on the same SSA loop, thus protecting their current investments!

Conclusions

Below is a summary of the benefits that the SSA technology brings to storage subsystems. SSA is a proven technology, designed specifically for high availability storage subsystems and is shipping in complete subsystems today. The main points to remember about SSA are:

1. Complete storage subsystems have been SHIPPING since August, 1995.
2. SSA is an OPEN industry standard embraced by computer industry manufacturers at all levels: chips, ASICs, adapters, peripherals, subsystems.
3. SSA is a LOW COST technology with single chip solutions available in off-the-shelf components.
4. SSA technology and design source is available to anyone in the industry, free of charge.
5. SSA provides HIGH PERFORMANCE with TRUE SCALABILITY to even higher system throughput.
6. SSA is the only storage interface with low enough cost for a PC environment and high enough performance for a mainframe environment.
7. SSA provides SUPERIOR FAULT ISOLATION and ERROR RECOVERY that was built into the architecture from day one.

Author: Michelle Tidwell, IBM Storage Systems Division, San Jose, CA.