White Paper

EMC²
where information lives

Architecture and Computer Storage

July 2002

Published in the USA

H749

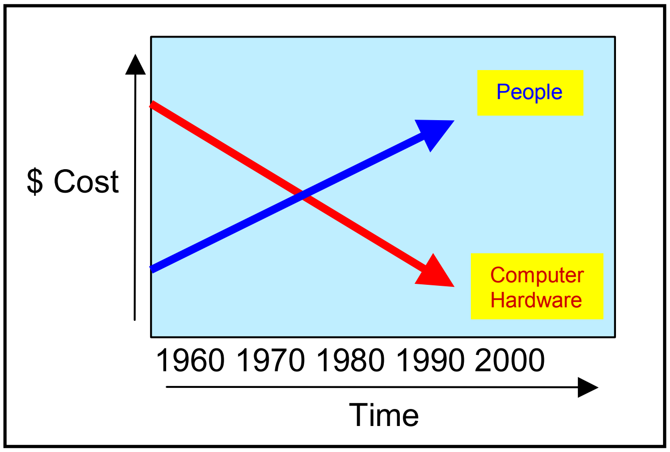
Table of Contents

Architecture and Computer Storage

The fifth definition of the word "architecture" in the Merriam-Webster dictionary is "the manner in which the components of a computer or computer system are organized and integrated".

Why is architecture important for computer systems?

For the entire history of computing, a fundamental economic trend has prevailed—computer hardware continues to become less expensive, and computer staff continues to become more expensive. As long as this trend continues, anything that helps to exploit new technology (at ever lower costs) while preserving the investment in people (at ever increasing costs to train, retain, and educate) is advantageous. A good computer architecture helps by being able to adapt to new technology without compromising the existing investment in people, software, and procedures.



In computer technology, a "perfect" architecture lasts forever, a "good" architecture typically lasts for more than a decade, and a "bad" architecture has a short useful life. Virtually all successful vendors in the computer industry have learned this lesson, some sooner than others. The concept didn't really present itself at all until the mid 1960s.

Architectures—Good and Bad

As it applies to computer technology, a "good" architecture tends to have some very specific benefits that differentiate it from a "bad" architecture:

- A good architecture preserves customer investments in existing software features and functions as it evolves.
- A good architecture preserves customer investments in training and staff knowledge as it evolves.
- A good architecture is known by the company it keeps—key third-party software and hardware suppliers design their products to integrate with products associated with a good architecture.
- A good architecture experiences multiple product updates over the duration of its life.
- A good architecture is able to exploit the leading edge of new technology on a continual basis.
- In contrast, there have tended to be some common characteristics of "bad" architecture:
- A bad architecture does not utilize key advances in technology, but rather is made obsolete by them.
- A bad architecture introduces alternative features that are incompatible with previous features that customers used and liked.
- A bad architecture is not well integrated with third-party software and hardware products.

Examples of architectures that have proven to be "good" abound in the systems and servers space. The greatest single attribute of a good architecture is often that it allows existing software to carry forward from one iteration of hardware to the next generation of hardware, thus preserving the customer's investment in people, training, procedures, and the software itself.

Probably the best example of "good" architecture among systems is the evolution of the IBM z/Architecture, most commonly known as the IBM Mainframe. When introduced in 1964 as the System/360, IBM broke from conventional thinking of the day by introducing a "family" of systems. Prior to this breakthrough, systems from all suppliers were usually unique and incompatible, even within their own company's offerings. The dramatic success of the System/360 concept became a model for others, as it was arguably the first systems architecture in the market. IBM has been able to keep this architecture alive to this day, as it evolved from the S/360 to the S/370, the S/390, and now finally the 64-bit zSeries.


IBM System/360 in 1965

Although they obviously understood the concept of systems architecture, IBM has had mixed success with architectures that have followed the System/360. IBM system architectures that have proven that they can evolve and survive have included the RS/6000 (now called the pSeries), and the AS/400 (now called the iSeries). In contrast, IBM has also had its share of system architectures that have failed to endure, including the Series/1, the 8100, and the System 34/System 36. Despite the lack of continuity, IBM's marketing efforts were no less aggressive for these failed architectures than for those that succeeded, right up until the day that each was cancelled.

IBM zSeries in 2002

Digital Equipment Corporation learned fast from IBM. In 1965, DEC introduced the PDP-8, which became their first architecture to have subsequent systems that were compatible. Unlike the large water-cooled, batch-oriented IBM mainframes of the day, the interactive PDP-8 revolutionized what became the mini-computer market. This architectural concept would continue to find dramatic success in Digital's PDP-11, VAX, and Alpha architectures. In the modern era of computing, Digital's only architecture that failed to survive the test of time was the DECsystem10/DECSYSTEM20, which abruptly ended its run in the early 1980s without warning. With Compaq's recently announced decision to move their OpenVMS and Tru64 operating systems off of the Alpha architecture and onto the Intel Itanium architecture, all remnants of Digital's CPU architectures are now effectively dead.

The concept of introducing systems within an architecture became the new de facto approach, and other successful architectures were introduced, such as the Sperry 1100/2200, the Data General Nova and Eclipse, the Hewlett-Packard HP3000 and HP9000, the Wang VS, Prime, Sun SPARC, and the Burroughs A and B series. When Sperry and Burroughs merged to become Unisys, they even took the concept further by integrating multiple architectures into their ClearPath systems.

The rise of Intel as a vendor-independent processor provider, and Microsoft as a vendor-independent operating system provider revolutionized the concept of architecture again. Intel's primary CPU architecture and Microsoft's MS-DOS, Windows, and Windows NT/Windows 2000 have evolved to become the de-facto standards for cross-vendor open platforms.

The server and O/S vendors that have survived into today's market "get it" – architecture is very important, and having a "good" architecture is mandatory to provide long-term differentiation in the market.

What about storage?

For decades, the only successful storage products were those that came as peripherals to a system or server purchase. Long before the term "server" became popular, it truly was a "system" that was implemented—if you bought an IBM CPU, then you bought IBM peripherals, if you bought a Sun CPU, then you bought Sun peripherals. By the late 1980's, large amounts of processing power had moved outside of the data center through the PC and mini-computer revolutions, but most of the data was still on monolithic systems in the data center. The storage systems that held the data virtually always had the same vendor logo as the CPU. Although most of the leading system vendors had created storage architectures that allowed for new technologies of disk drives to support previous storage controllers, they typically used unique and proprietary interfaces to their disk subsystems that precluded anyone else from connecting. Most system vendors would even wage or threaten legal battles against anyone who tried to reverse engineer and connect to these proprietary interfaces.

Some independent storage companies such as CDC, Memorex, Systems Industries, MTI, and others attempted to enter the storage market against the system vendors with storage-only offerings during the 1970's and 1980's, but none were able to sustain enduring success. Possibly the greatest reason for their failure was that these companies tended to compete with a basic strategy that included claiming roughly similar capabilities to the system vendor offerings, but at lower prices. Without any dramatic functional advantages, they were vulnerable to concerns regarding their ability to continually integrate into new systems technology. Customers would almost always challenge the service and support capabilities of these smaller vendors when compared against the size, reputation, experience, and integration capabilities of the server vendors, with whom they were much more familiar.

Breakthrough #1 that changed everything about storage – RAID

Using some form of Redundant Array of Independent Disks (RAID) technology has become so ubiquitous for storage with most production computing environments, that it almost seems hard to believe that until around 1990 virtually all storage involved unprotected, physically large single disk drives. Patterson, Gibson, and Katz developed the concept of RAID at the University of California at Berkeley in the late 1980s. An excellent source of background on RAID is available from the RAID Advisory Board at: http://www.raid-advisory.com/rabguide.html.

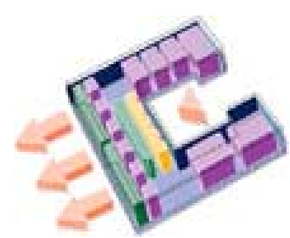
While virtually all vendors that survived in storage through the 1990's immediately picked up on the importance of RAID , only three succeeded in quickly delivering products to exploit it. One was Digital Equipment Corporation, who in late 1989 announced a proprietary software implementation of RAID-1 (volume shadowing) and RAID-0 (striping) for their high-end VAX/VMS Cluster systems. The other two were EMC Corporation and Data General Corporation.

Breakthrough #2 that changed everything about storage – EMC's MOSAIC storage architecture and Integrated Cached Disk Array (ICDA)

EMC Corporation began in 1979 as a memory supplier for systems made by IBM, DEC, Prime, Wang, and HP. EMC would not ultimately find success in the add-in memory market; however, core competencies in advanced memory technologies would later prove overwhelmingly advantageous. EMC also benefited (through necessity in as much as design) in multi-vendor/multi-platform engineering and support.

By the late 1980's EMC had leveraged its knowledge in memory to build solid-state disk devices, which basically appeared to an operating system as if they were a physical disk drive, but in reality were built upon radically higher performing memory boards. The problem with solid-state disks was volatility—if power was lost, so was data—unlike conventional disks. EMC experimented with several ways to solve this, and out of those efforts came a whole new approach to storage from a senior engineer named Moshe Yanai. This new category of product was called an Integrated Cached Disk Array, or ICDA.

The concept of an ICDA is simple, but the intelligence in software required to implement it successfully is anything but simple. In an ICDA, the entire storage subsystem would appear to a computer server as merely a set of conventional disk drives. To prevent delays caused by physically involving the mechanical nature of disks through reads and writes to disk, the ICDA would use multiple microprocessors and cache memory to provide drastically higher performance without any impact to the computer server.



The key to successfully delivering EMC's ICDA was the foresight to design the first product within the framework of an architecture, rather than to merely build a single point product. The temptation to rush a point product with such an innovative approach to market was immense, but the discipline to design the architecture foremost opportunely prevailed. The architecture was called MOSAIC:2000. The "MOSAIC" portion of the name was to describe the concept of being able to change any "tile" in the architecture with new technology without disrupting the other "tiles" in the picture. The "2000" portion of the name was to imply that the architecture should be able to survive well into the next millennium.

By the design of MOSAIC: 2000, inside the ICDA there would be front-end microprocessors that would take all read and write requests and service them to and from cache memory at drastically higher speeds than were possible with conventional disks. All I/Os would be serviced from cache memory. There would also be back-end microprocessors that would independently service all I/Os between cache and the physical storage devices. Multiple high-speed busses would provide interconnectivity between the front-end microprocessors and cache, and the back-end microprocessors and cache.

Symmetrix in 1991

A fundamental aspect of the MOSAIC: 2000 architecture was that it was expected that advances in technology would always be incorporated into new product releases as they became viable. Entire classes of technology could change, yet the architecture should be designed to endure. The product that emerged from within the architecture was Symmetrix—introduced in 1990.

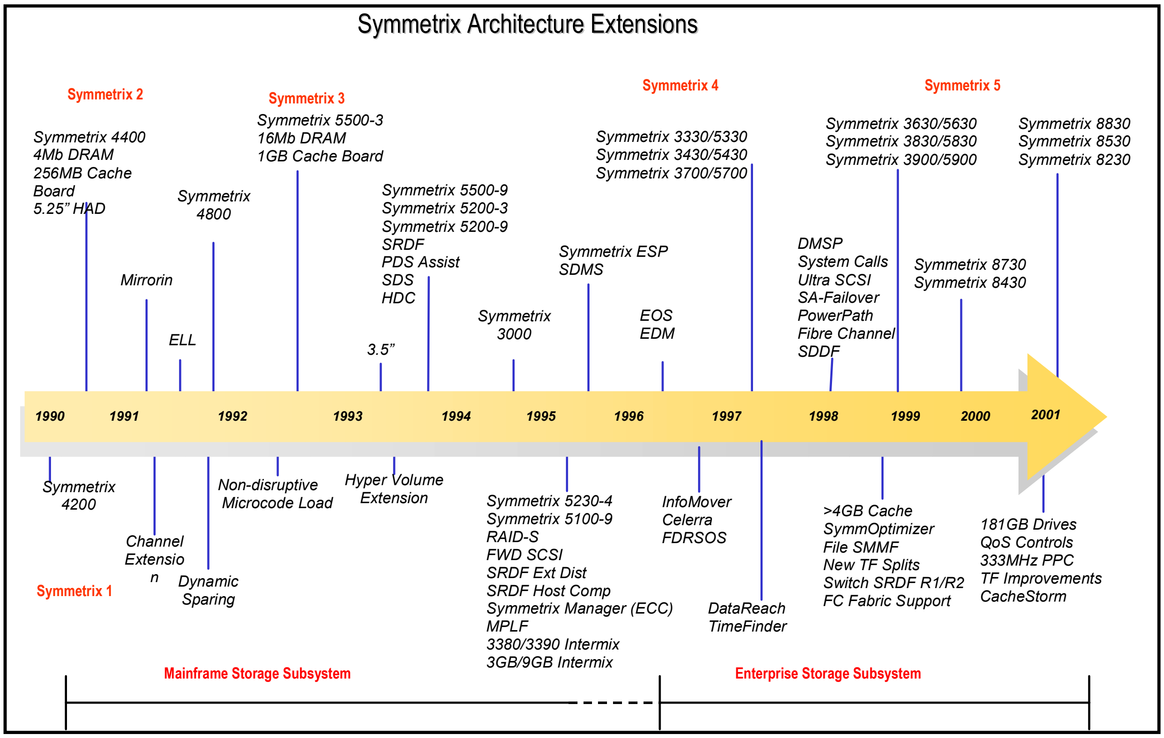
Symmetrix was a fitting name for a product within this architecture for several reasons. There were no single points of failure in the hardware design. Busses were symmetric to each other, and fault-tolerant. The front-end microprocessors were located on cards called Channel Directors that were symmetric to the back-end microprocessor cards, called Disk Directors. Cache boards were symmetric to each other, as were power supplies and fans. The core software that ran Symmetrix (later to be named Enginuity) was the industry's first storage-specific operating system. The unique algorithms and operations of Enginuity would continue to be the greatest technical differentiating factor between Symmetrix and all other storage offerings.

Symmetrix in 2002

The first Symmetrix supported IBM mainframe parallel channel (bus and tag) interconnection, but unlike alternative DASD offerings, Symmetrix was the first to offer a true RAID implementation in all aspects – **R**edundancy rather than unprotected disks (with RAID-1 protection), and a true **A**rray of **I**ndependent **D**isks rather than single disks.

While it was the unique performance and availability features of Symmetrix that made it so popular in the early 1990s, its ongoing success can be attributed to having been designed within an architecture. Symmetrix has consistently demonstrated the value of the MOSAIC architecture by being able to adapt and incorporate virtually every important new storage technology that has emerged since 1990, and is now in its fifth generation of platforms. Some of these technologies have included:

- Staying on the absolute leading edge of new disk technology: 5 ¼” 3GB, 9GB, 23GB, 47GB and 3 ½” 4GB, 9GB, 18GB, 36GB, 73GB, 181GB
- Front end connectivity: All industry standard and de-facto standard interconnects, including Mainframe Parallel Channels (bus and tag), ESCON, FWD SCSI, Ultra SCSI, FC-AL, FC-SW, FICON, IP, multi-site with T1, T3, ATM, FDDI, FC, ESCON, IP
- Back-end connectivity: SCSI, FWD SCSI, Ultra SCSI, LVD SCSI
- New cache technology: The world's highest capacity (64GB), and the world's only 16-way cache
- Microprocessors: From a few Motorola 68000s, now up to (80) 333MHz PowerPCs (the world's greatest processing power for a storage product)
- From 2 to 4 system busses with consistently increasing bandwidth, currently at 400MB/s each (the world's fastest internal storage interconnect paths)
- Support for over 40 different operating systems, all major cluster implementations, and all major platforms (the world's broadest and deepest support, by a very wide margin)



During the mid to late 1990s, storage software became one of the most significant factors to differentiate storage platforms, and in particular to accelerate the success of Symmetrix. Due to the architectural design of Symmetrix under MOSAIC, EMC was able to add the richest suite of software functions in the industry to existing and future products within the existing architecture. These functions were all added upon the stable base of Enginuity. Some of these products, such as Symmetrix Remote Data Facility (SRDF) remote data replication software and TimeFinder local data replication software, have gone on to become de-facto industry standards for business continuance solutions. *"If any technology product emerged the hero from the recent disasters, it was an EMC product called SRDF."* – Barron's, September 24th, 2001.

As EMC continued to introduce new software products and enhancements to existing software products that ran with Symmetrix, other third party software companies became interested in integrating with EMC storage platforms at both the hardware and software level. Proving the flexibility of the architecture once more, EMC introduced the EMC Developers Program, which allows program members to utilize EMC's open Application Programming Interfaces (APIs) to directly access Symmetrix features and functions. This provides the ability of a third-party software vendor to enhance their own products with specific "hooks" into EMC product functions. Today, the EMC Developers Program includes just over 100 companies as members, who have collectively delivered 66 generally available software products that use EMC's open APIs. This level of integration tends to provide incremental functionality for the offerings of these independent software vendors that is available when used with EMC's technology, but is unavailable within their products when used with alternative vendor storage products.

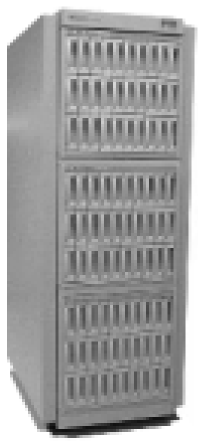
An aspect of Symmetrix and MOSAIC that continues to set it apart from all alternative storage offerings is the ability of older Symmetrix hardware to support the most current versions of Enginuity and related software offerings. This provides functional investment protection, and dramatically extends the useful life of Symmetrix hardware platforms as viable solutions.

Breakthrough #3 that changed everything about storage: CLARiiON—Data General's modular open storage.

Simultaneous with the development of Symmetrix at EMC, Data General Corporation was building a very different storage architecture that had its roots in a fault-tolerant computer system first introduced in 1990. While the DG fault-tolerant processor portion of the system never did make it to market, the storage component of the system did. The design team was led by senior engineer Bob Solomon and team manager Paul Suffredini. It was originally called the High Availability Disk Array (HADA), and unlike all other storage systems that were available for UNIX or Windows servers it used a redundant, fault-tolerant architectural design.

The concept of "open systems" was certainly not new to the UNIX and Windows systems environments. Despite efforts to provide software portability between different UNIX offerings, and until Data General's efforts to provide a single storage system that could run with different UNIX or Windows systems, the server vendor had proprietarily owned virtually all storage for "open systems."

In 1992, the CLARiiON business unit was formed to provide the industry's first "open" storage for "open systems". CLARiiON stood out from all other products in the open systems space not only because of its potential to work with different servers from different vendors, but because of the fundamental core advantages of its architecture.



Early CLARiiON

Like the Symmetrix, CLARiiON products were built within the framework of an architectural design. The basic building blocks of the CLARiiON were two separate components – Control components and Disk components that connected to each other through industry-standard buses. These evolved into the Disk Processor Enclosure (DPE) and the Disk Array Enclosure (DAE). The DPE contains the redundant microprocessors, mirrored cache memory, and interconnectivity components of a CLARiiON, as well as redundant power, redundant cooling, and redundant access to up to 10 disk drives. DAEs connect to a DPE, and provide incremental disk capacity (up to 10 disks per DAE) with redundant connectivity, power, and cooling. Both the DPE and DAE are designed within the architecture to allow for independent technology enhancements, while preserving the investment in the existing components and software.

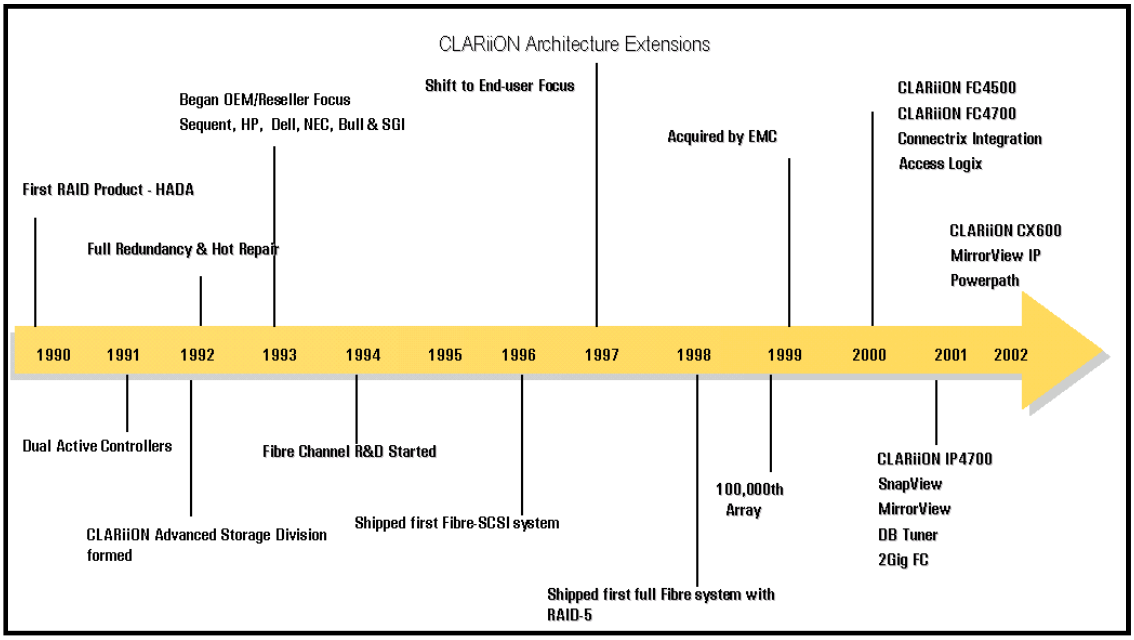
CLARiiON in 2002

Unlike EMC, which initially sold the Symmetrix through a direct sales force to the mainframe market, Data General targeted the open systems OEM channel for UNIX and Windows servers. By 1994, Data General had successfully signed on StorageTek, Hewlett-Packard, NEC, Bull, and SGI to resell CLARiiON as their premier storage offerings. As both the "best in class" open storage product in the market, and by virtue of being the first true "open" storage product, CLARiiON became very successful.

Over time, CLARiiON would see major technology enhancements to the architecture, while preserving existing customer investments. Some of these technologies included:

- Staying on the absolute leading edge of new disk technology: 3 ½" 4GB, 9GB, 18GB, 36GB, 73GB
- Front-end connectivity evolution from SCSI to FC-AL and FC-SW, including 2 gigabit technology
- Back-end connectivity evolution from SCSI to Fibre Channel
- Microprocessor evolution from Motorola to (4) Intel Pentium CPUs at 2GHz
- Support for all major UNIX and PC operating systems
- All major RAID options: RAID 0, 1, 0+1, 3, and 5
- Unique ability to participate in either a SAN or NAS environment by simply specifying the appropriate front-end cards, or to change from one environment to the other.
- Appending 8 extra bits to each block, providing inherent on-disk data integrity features that far surpass alternative offerings

- SNiiFER software that constantly monitors system integrity
- Destage of write cache to preserve data integrity during extended power outages



CLARiiON has even been able to allow for new upgraded models of the DPE to work with existing installed DAEs.

Software for CLARiiON has become a major differentiating factor, and has developed to where all major Enterprise-class storage software functions are available today:

- Local and remote data replication
- Path failover
- SAN LUN masking
- Database tuning
- Extensive operational and performance management

Summary

> *"He who does not lay his foundation beforehand may, by great abilities, do so afterwards - although with great trouble and cost to the architect and great danger to the building!"*—Machiavelli—*The Prince*—1513

Good architectural design in computer related products results in significant long-term customer economic value and improved productivity. Merely adapting the latest technology at the expense of good architecture compromises those benefits. The ability to deliver an enduring architecture has proven to be possibly the greatest factor that impacts vendor success in the market. Within the same timeframe that CLARiiON and Symmetrix have endured, virtually every alternative storage vendor has introduced and withdrawn at least two storage architectures, and most have introduced and withdrawn many more. To date, only these two storage architectures have proven their ability to deliver those benefits—CLARiiON and Symmetrix.