INTRODUCTION TO EMC XTREMSF

- XtremSF is server-based PCIe Flash hardware
- XtremSF can be used as local storage or as a caching device with EMC XtremSW Cache

Abstract

This white paper provides an introduction to EMC XtremSF. It describes the PCIe flash hardware and provides details on implementation, usage, performance, and benefits.

May 2013
# Table of Contents

Executive summary .................................................................................................................. 4

Introduction .............................................................................................................................. 5
  Audience .............................................................................................................................. 5
  Terminology .......................................................................................................................... 5

Use cases of Flash technology ................................................................................................. 6
  Advantages of XtremSF. ......................................................................................................... 6
  Flash cell architecture .......................................................................................................... 7

XtremSF design concepts ......................................................................................................... 10
  Business benefits .................................................................................................................. 12

Implementation details ........................................................................................................... 13
  XtremSF as DAS ................................................................................................................... 13
  Data protection with DAS...................................................................................................... 14
  XtremSF with XtremSW Cache ........................................................................................... 15
  XtremSF with XtremSW Cache in split card mode ............................................................... 15

Performance Considerations .................................................................................................. 17
  Workload characteristics ...................................................................................................... 17
  Throughput versus latency ................................................................................................... 18
  XtremSW Cache in split card mode ..................................................................................... 18
  Other bottlenecks in the environment .................................................................................. 18

Usage guidelines and characteristics ...................................................................................... 19
  Specifications ....................................................................................................................... 20
  Constraints .......................................................................................................................... 20

Application use cases and performance .................................................................................. 21
  Test results .......................................................................................................................... 21

Conclusion ............................................................................................................................... 23

References ............................................................................................................................... 24
Executive summary

Since the first deployment of Flash technology in disk modules (commonly known as solid-state drives or SSDs) by EMC in enterprise arrays, EMC’s goal has been to expand the use of this technology throughout the storage environment. Due to the decreasing costs of Flash technology and requirements for high application performance, server-based Flash has become an integral part of this expansion.

EMC XtremSF is server-based PCIe Flash hardware that reduces latency and increases throughput to dramatically increase application performance. XtremSF can be used as a direct-attached storage (DAS) device or as a caching device in conjunction with EMC XtremSW Cache server flash caching software. When used as DAS, data sets are stored locally for accelerated reads and writes. When used with XtremSW Cache, the intelligent caching algorithm accelerates reads, while all writes persist to the networked storage for high availability, integrity, and disaster recovery. This white paper will primarily focus on the DAS use cases of XtremSF. For more details on the caching use case, see the Introduction to EMC XtremSW Cache white paper available on www.EMC.com.

Implemented either as DAS or as a caching tier with XtremSW Cache software, XtremSF is an excellent solution for applications with low latency and high I/O requirements. XtremSF is optimal for high-transactional and/or high-performance workloads often associated with web 2.0 applications, virtual desktop infrastructure (VDI) environments, high-performance computing (HPC), and high-performance trading applications.

XtremSF benefits include:

- When used as a local storage device, XtremSF provides performance acceleration for both read and write-intensive workloads.
- With its 2.2TB half-height, half-length (HHHL) eMLC card, XtremSF provides the industry’s largest capacity in the smallest PCIe form factor.
- XtremSF offers the choice of enterprise-grade multi-level cell (eMLC) or single-level cell (SLC) Flash for performance, capacity, and price-optimized solutions.
- XtremSF provides high levels of concurrency for applications that require multiple I/Os to be processed in parallel.
- When used with XtremSW Cache, XtremSF enables accelerated performance with the protection of the back-end network storage array.
Introduction
This white paper provides an introduction to XtremSF. Topics covered in this white paper include implementation in physical environments, performance considerations, best practices, usage guidelines, characteristics, and some application-specific use cases.

Audience
This white paper is intended for organizations who are considering the use of XtremSF in their storage environments. It assumes a basic understanding of Flash technology and its benefits.

Terminology
Dataset: Span of data being managed by an application; for example, the database size
Working set: The amount of data actively accessed within the dataset
Queue depth: The number of pending I/O requests from the application
Use cases of Flash technology

There are different ways in which Flash technology can be used in an environment depending on the use case, application, and customer requirements. EMC’s architectural approach is to use the right technology in the right place at the right time. This includes using Flash:

- As direct-attached storage
- As a cache in a server
- As a cache in an array
- As a storage tier in an array
- In an all-Flash array

There are different types of Flash with different cost structures, endurance considerations, and performance characteristics. All types of Flash have a proper place in the vast continuum of use cases. Some of the use cases for Flash are below:

- Applications with performance requirements with or without protection requirements that are read and write heavy may be a good fit for PCIe Flash in the server as DAS—for example, XtremSF.
- Applications with high performance and protection requirements that are read heavy are a perfect fit for PCIe Flash in the server as a cache—for example, XtremSF combined with XtremSW Cache.
- Applications with performance and protection requirements that are read and write heavy may be a good fit for Flash in the array as a cache—for example, EMC FAST Cache on an EMC VNX storage system.
- Applications with mixed workloads and changing data “temperatures” are a perfect fit for Flash as part of a tiering strategy—for example, Fully Automated Storage Tiering for Virtual Pools (FAST VP) on an EMC VMAX storage system.
- Applications requiring highly consistent performance may be a good fit for an all-Flash array—for example, EMC XtremIO.

Advantages of XtremSF

There are several advantages to using XtremSF hardware for application acceleration:

- XtremSF’s industry-leading performance reduces latency and increases throughput to dramatically improve performance in read and write-heavy applications.
- The large capacities offered in the XtremSF portfolio allow users to store large data sets or multiple working sets on a single PCIe Flash device.
- The range of capacity options within the XtremSF line-up provides the flexibility of adding the exact amount of Flash needed to accelerate target applications.
- The choice between eMLC and SLC NAND Flash offers customers the flexibility of choosing the Flash memory that suits their individual needs. (See Flash cell architecture below for descriptions of eMLC and SLC.)
- XtremSF provides simple installation and management features for an easy, customer-installable solution.

**Flash cell architecture**

In general, there are two major NAND-based Flash cell technologies used in all Flash drives:

- Single-level cell (SLC)
- Multi-level cell (MLC)

A cell is the smallest unit of storage in any Flash technology and is used to hold a certain amount of electronic charge. The amount of this charge is used to store binary information.

NAND Flash cells have a very compact architecture; their cell size is almost half the size of a comparable NOR Flash cell. This characteristic, when combined with a simpler production process, enables a NAND Flash cell to offer higher densities with more memory on a given semiconductor die size. This results in a lower cost per gigabyte.

Flash storage devices store information in a collection of Flash cells made from floating gate transistors. SLC devices store only one bit of information in each Flash cell (binary), whereas MLC devices store more than one bit per Flash cell by choosing between multiple levels of electrical charge to apply to its floating gates in the transistors (See Figure 1).
Since each cell in MLC Flash has more information bits, an MLC Flash-based storage device offers increased storage density compared to an SLC Flash-based version. However, MLC Flash has lower performance and endurance because of its inherent architectural tradeoffs. Higher functionality further complicates the use of MLC Flash, which requires advanced Flash management algorithms and controllers. There are two grades of MLC Flash produced today: consumer-grade (cMLC) used in consumer storage products such as thumb drives, and higher quality enterprise-grade (eMLC) used in the MLC versions of XtremSF.

SLC Flash and MLC Flash offer capabilities that serve two very different types of applications—those requiring high performance at an attractive cost per bit (eMLC), and those who are less cost-sensitive and seek even higher performance and endurance over time (SLC).

Taking into account the varying types of I/O profiles and requirements of enterprise applications, EMC XtremSF provides customers the flexibility of choosing between eMLC and SLC Flash architectures.

Table 1 compares the SLC and MLC Flash characteristics (typical values).

---

Table 1: SLC and MLC Flash comparison

Although SLC Flash offers a lower density, it also provides an enhanced level of performance in the form of faster reads and writes. Because SLC Flash stores only one bit per cell, the need for error correction is reduced. SLC also allows for higher write/erase cycle endurance, making it a better fit for use in applications requiring increased endurance and viability in multi-year product life cycles.

For more details on various Flash cell architectures, refer to the *Considerations for Choosing SLC versus MLC Flash Technical Note* on the EMC Support website (https://support.emc.com).
**XtremSF design concepts**

Over the past decade, server processing technology has continued to advance along the Moore’s Law curve. Every 18 months, memory and processing power have doubled, but disk drive technology has not. Spinning drives continue to spin at the same rate. This has caused a bottleneck in the I/O stack whereby the server and the application have capacity to process more I/O than the disk drives can deliver. This is referred to as the I/O gap, as shown in Figure 2.

![Figure 2: I/O gap between the processor and storage sub-systems](image)

Flash drives in the storage system have helped to close this gap, as Flash is a silicon technology—not mechanical—and can therefore enjoy the same Moore’s Law curve. Figure 3 shows a comparison of different storage technologies based on the I/Os per second (IOPS) per gigabyte (GB) of storage that they offer.
Mechanical spinning drives provide a great dollar-per-gigabyte economic value to cold datasets, but they do not provide the best performance. Putting SSDs in the array provides an order-of-magnitude better performance. Putting PCIe Flash in the server can accelerate performance by yet another order of magnitude over SSDs.

Placing key data sets on PCIe Flash in the server can decrease the latencies associated with local disks, resulting in an increase of performance by up to 40 percent in enterprise applications.

XtremSF ensures the accessibility of data by storing it microseconds away on the server’s PCIe bus. Figure 4 shows an XtremSF deployment in a typical DAS environment.

**Figure 3: Comparison of storage technologies**

![Comparison of storage technologies](image)

**Figure 4: Typical XtremSF deployment**

![Typical XtremSF deployment](image)
XtremSF is designed to provide:

- **Performance**: Reduces latency and increases throughput to dramatically improve application performance
- **Capacity**: Allows users to install as much or as little PCIe Flash capacity as needed for their specific application requirements
- **Flexibility**: Provides options of eMLC and SLC Flash technology to meet specific endurance and performance needs

**Business benefits**

XtremSF provides the following business benefits:

- XtremSF increases the performance and reduces the response time of applications. For most businesses, this translates into an ability to do faster transactions, and more of them.
  - For example, a company using an Oracle or SQL Server OLTP database may be limited in the number of transactions it can process because of the number of IOPS that the local drives can provide. XtremSF increases throughput to allow for more transactions, thereby generating more revenue for the company.
- When used with XtremSW Cache, customers get the following additional benefits:
  - Typical customer environments might have multiple applications accessing the same storage system on the back end. Some of these applications are more important than others. Users want to get the best performance for these applications while making sure that the other non-critical applications continue to get “good enough” performance.
  - XtremSF provides flexibility because it is installed in the server instead of in the storage network. With multiple applications accessing the same storage, XtremSF improves the performance of the application on the server where it is installed, while other applications on other servers continue to get good performance from the storage system. In fact, they might get a small performance boost because part of the back-end storage system’s workload gets offloaded to XtremSF, and the storage system has more processing power available for these applications.
- XtremSF is infrastructure agnostic. It can accelerate any type of application running a wide variety of operating systems.
- Split card mode allows you to use part of the server Flash for cache (with EMC XtremSW Cache) and the other part as DAS.
Implementation details

This section of the white paper provides details about how I/O operations are handled when XtremSF is installed on the server. In a typical DAS implementation of XtremSF, the following components need to be installed in the environment:

- Physical XtremSF device
- XtremSF device driver

For more information about the installation of these components, see the *EMC XtremSF Installation and Administration Guide.*

If XtremSF will also be used for data caching with a back-end storage array, the following must also be installed in the environment:

- XtremSW Cache software

For more information about the installation of these components, see the *EMC XtremSW Cache Installation and Administration Guide.*

Figure 5 shows a simplified form of the XtremSF DAS architecture. The server consists of two components—the application layer and the XtremSF device in the server.

XtremSF hardware is inserted in a PCIe Gen2 x8 slot in the server and the driver is installed at the operating system level. Once installed, the XtremSF card is configured with a file system and partitioned using the operating system’s logical volume manager. The target application is then configured to read from and write to a particular volume on XtremSF. If needed, the entire XtremSF card can also be configured as a single volume.

In the following examples, XtremSF is configured as DAS for storing data locally and as a cache when combined with XtremSW Cache. Based on the solution requirements, the best suited configuration can be used to gain the desired performance benefit.

**XtremSF as DAS**

In this example, the XtremSF card is being used for local data storage. For a DAS use case, eMLC Flash-based PCIe cards are adequate from a performance and endurance requirement perspective. Figure 5 below shows the details of a DAS use case.

1. The application writes data to the XtremSF card.
2. When needed, the application reads the data from XtremSF.

This use case provides significant throughput and latency benefits to the application.
Data protection with DAS

XtremSF as DAS benefits both temporary and business-critical data. However, since PCIe Flash does not provide the data protection benefits inherently found on back-end storage arrays, business-critical data must be protected in other ways. The best way to do this is to utilize data protection features at the operating system or application level.

Native operating system tools such as logical volume managers (LVMs) can combine multiple storage devices (as seen by the OS) into RAID groups to provide performance (RAID 0) or protection (RAID 1) for application data. With multiple XtremSF devices installed in a single server, LVMs can be used to create RAID Groups and optionally provide performance and/or protection.

Application-level replication services are provided by a number of database and enterprise applications. Services such as Oracle Data Guard, Oracle Real Application Clusters (RAC), Oracle Automatic Storage Management (ASM) and Microsoft SQL Servers AlwaysOn can provide the high availability and data replication protection required. These are just a few examples of how data protection can be applied to XtremSF. Each installation will have a specific solution best suited for the applications and operating systems in use.

Host-based replication tools can provide replication functionality for locally stored data. For example, there are block level replication software tools that provide real-time continuous replication of any data between servers across a LAN or a WAN. These are often implemented as an operating system filter driver that sits on the file system in the I/O stack. Some of these tools can provide both synchronous and asynchronous replication.
**XtremSF with XtremSW Cache**

XtremSF hardware can also be used in conjunction with XtremSW Cache software to create a caching solution for improved application performance while protecting data on a back-end storage array. Figure 6 shows the basic configuration of XtremSW Cache installed with XtremSF.

**Figure 6: XtremSF with XtremSW Cache**

In this configuration, a copy of the application’s “hottest” data is stored on the local PCIe card for read acceleration, while writes are persisted to the storage array. For a detailed description of an XtremSW Cache deployment with XtremSF, see the *Introduction to EMC XtremSW Cache* White Paper.

**XtremSF with XtremSW Cache in split card mode**

EMC XtremSW Cache software has a unique split card feature, which allows users to use a portion of the server Flash as a caching device with XtremSW Cache, and the remainder as DAS. This feature provides the ability to simultaneously combine the two previously described use cases.

Figure 7 shows a representation of an XtremSF card used in split card mode.
As is the case with a 100 percent DAS use case, the contents of the DAS portion do not persist to any storage array. Therefore, it is recommended that the user store temporary data on the DAS portion, or utilize an application or operating system level data protection tool.
Performance Considerations

As DAS, XtremSF is a sophisticated Flash memory storage solution, so there are certain considerations that should be taken into account when evaluating XtremSF performance.

Workload characteristics

The final performance benefit that you can expect from XtremSF depends on the application workload characteristics. EMC recommends that you do not use XtremSF as local storage for applications that do not have a suitable workload profile. For example:

- **Data set size:** You should have an idea of the data set size of the application relative to the XtremSF card capacity. The entire data set can be placed onto the XtremSF card for maximum performance benefit.

- **Random versus sequential workloads:** An EMC storage array is very efficient in processing sequential workloads from your applications. The storage array uses its own cache and other mechanisms like “prefetching” to accomplish this. However, if there is any randomness in the workload pattern, the performance is lower because of the seek times involved with accessing data on mechanical drives. The storage array cache is also of limited use in this case because different applications using the storage array will compete for the same storage array cache resource. Flash technology does not have any latency associated with seek times to access the data. XtremSF will therefore show maximum performance improvement when the application workload has a high degree of randomness.

- **Concurrency:** Mechanical drives in the storage array have only one or two read/write heads, which means that only a limited number of I/Os can be processed at any one point in time from one disk. So when there are multiple threads in the application trying to access data from the storage array, response times tend to go up because the I/Os need to wait in the queue before they are processed. However, storage and caching devices using Flash technology typically have multiple channels internally that can process multiple I/Os at the same time. Therefore, XtremSF shows the maximum performance difference when the application workload has a high degree of concurrency. The application should request multiple I/Os at the same time.

- **I/O Size:** Large I/O sizes tend to be bandwidth-driven and reduce the performance gap between Flash technology and non-Flash technologies. Although XtremSF will enhance performance over a wide range of I/O sizes, applications with smaller I/O sizes (for example, 4KB or 8 KB) will show the maximum performance benefit.
**Throughput versus latency**

There are some applications that can push the storage environment to the limit to provide as many IOPS as possible. Using XtremSF in those application environments will show very high IOPS at very low response times. However, there are also applications that do not require very high IOPS, but they require very low response times. You can see the benefit of using XtremSF in these application environments as well. Even though the application issues relatively few I/Os, whenever the I/Os are issued they will be serviced with a very low response time. For example, a web application may not have a lot of activity in general, but whenever a user issues a request, the response will be very quick.

**XtremSW Cache in split card mode**

When this functionality is used, the same Flash resources are shared between the cache and DAS portions. Therefore, the cache performance may be less compared to when the PCIe card is being used solely as a caching device.

**Other bottlenecks in the environment**

By making data even more accessible to the application, XtremSF helps improve throughput and reduce latencies. However, any drastic improvement in application throughput may expose new underlying performance bottlenecks and/or anomalies within the hardware or software stack.
Usage guidelines and characteristics

This section provides some of the usage guidelines and salient features of XtremSF.

- XtremSF is optimized for highly concurrent workloads, providing increased performance with higher thread counts. This design takes advantage of Flash media’s ability to support high bandwidth rates. By accessing fewer Flash chips for each I/O, chip level bottlenecks that typically occur with multiple threads are avoided. Because of this, applications that can utilize multiple threads will yield the greatest performance benefit.

- XtremSF has fixed mapping unit sizes that are optimized to match the larger I/O sizes (4K or 8K) that are typically associated with the I/O sizes of common enterprise applications.

- On most XtremSF models, Flash management is done within the card, offloading these tasks from the host server. This reduces CPU utilization and the amount of DRAM required in the host.

- Data integrity is a key concern. It is critical to ensure the data returned on a read mirrors what was last written to the device. To provide data integrity on XtremSF, an end-to-end checksum is used to verify the data read exactly matches the data previously written to the block.

- To protect the Flash from failure, XtremSF cards use a RAID-parity scheme, where each element in the RAID group is processed by a different Flash channel.

- In the event of a failure of a Flash chip on the PCIe card, XtremSF automatically detects and quickly rebuilds the data.

- Garbage collection and wear leveling are performed with minor impact to performance. These activities are performed at the Flash chip level, and take advantage of XtremSF’s high concurrency capability, reducing the impact to the application.

- When used with XtremSW Cache as a caching device, writes are synchronized with the array. Write operations issued by the application will be limited by the speed with which the back-end array can process the writes. When used in split card mode, this only affects the caching portion of the card; the DAS portion is unaffected by the array's performance.

- When XtremSF is used as a storage device, both reads and writes are accelerated. To protect business-critical data stored on the card, application or operating level data protection features should be used.
Specifications

- XtremSF is optimized for 4K and 8K I/O workloads, but will work seamlessly and provide benefit with applications with other predominant I/O sizes.

- XtremSF needs to be installed in a PCIe Gen2 x8 slots in a rackmount server. It can be installed in x16 PCIe slots also, but only 8 channels will be used by XtremSF. Similarly, if it is installed in an x4 PCIe slot in the server, XtremSF will perform sub-optimally.

- XtremSF is designed to maintain a <25W power requirement per the PCIe 2.0 specification.

- XtremSF cards are available in the following capacities:
  - Single-Level Cell (SLC)
    - 350 GB and 700 GB
  - Multi-Level Cell (eMLC)
    - 550 GB, 700 GB, 1.4 TB and 2.2 TB

- Multiple XtremSF cards can be used per server. A single logical source volume can be created across multiple cards by using an operating system’s LVM.

- Multiple source volumes can be created on a single card using an operating system’s LVM.

- XtremSF is compliant with the Trade Agreements Act (TAA). The following main requirements are certified as not applicable to XtremSF:
  - FIPS 140-2
  - Common Criteria
  - Internet Protocol version 6

Constraints

- XtremSF provides Flash-level data integrity protection by using a RAID-parity scheme. However, XtremSF does not provide data protection services for business-critical data. If data protection is required, XtremSW Cache or other software tools must be used.

- Blade servers require a customized version of the card and therefore do not support XtremSF. However, servers that enable a PCIe expansion card to be connected to their chassis can work with XtremSF cards. For the most current list of supported operating systems and servers, refer to E-Lab Interoperability Navigator.
Application use cases and performance

XtremSF is optimal for high-transactional and/or high-performance workloads often associated with web 2.0 applications, virtual desktop infrastructure (VDI) environments, high-performance computing (HPC), and high-performance trading applications. It can also be used to accelerate analytics, reporting, data modeling, indexes, database dumps, batch processing, background tasks, and other temporary workloads.

When using XtremSF as DAS, the greatest benefit will be achieved in high read and write applications with low latency requirements. If XtremSF is combined with XtremSW Cache for caching, you will achieve the greatest results in high-read applications and applications with a highly concentrated skew of data.

Test results

EMC conducted application-specific tests with XtremSF to determine potential performance benefits when this product is used. Here is a summary of the XtremSF benefits with some popular applications:

SQL Server

With a TPC-E like workload in an 800 GB Microsoft SQL Server 2012 database environment on a Cisco UCS server, XtremSF delivered thirty times more IOPS than the same workload running on multiple local disks. Figure 10 shows this performance improvement.

![SQL Server Database Performance](image)

*Figure 8: XtremSF performance vs. local disk with a SQL Server database*
Oracle

With a TPC-E-like OLTP workload in a 1.2 TB Oracle 11g R2 physical environment, XtremSF recorded forty times the performance when compared to the same workload running on multiple local disks. Figure 11 shows this performance improvement.

![Oracle Database Performance Graph](image)

Figure 9: XtremSF performance vs. local disk for an Oracle Database

For more information on application-specific guidelines and test results, refer to the list of white papers provided in the References section.
Conclusion

There are multiple ways in which Flash technology can be used in an environment—in the server or the storage array, as a cache or a tier. The key, however, is matching the hardware and software technology with the application for the maximum performance benefit.

With its industry-leading performance, XtremSF dramatically accelerates the performance of both read and write-intensive applications. XtremSF is a flexible hardware platform that offers a range of capacity points, accommodating small to large datasets. Cards are available in both SLC and eMLC Flash, offering greater flexibility when fulfilling specific performance and endurance requirements. XtremSF provides the densest Flash server storage in the smallest PCIe form factor with its 2.2TB HHHL card. XtremSF can be used alone as DAS or as a cache when combined with XtremSW Cache.
References

The following documents are available on https://support.emc.com:

- Analyst White Paper: Demartek—EMC XtremSW Cache Flash Caching Solution Evaluation
- Analyst White Paper: ESG—EMC’s Flash Strategy
- Data Sheet: EMC XtremSF
- Data Sheet: EMC XtremSW Cache
- Installation and Administration Guide: XtremSF for Windows and Linux
- Installation and Administration Guide: XtremSW Cache for VMware
- Installation and Administration Guide: XtremSW Cache for Windows and Linux
- Installation and Administration Guide: XtremSW Cache VMware Plug-in
- Release Notes: XtremSF for Windows and Linux
- Release Notes: XtremSW Cache for VMware
- Release Notes: XtremSW Cache for Windows and Linux
- Technical Note: Considerations for Choosing SLC versus MLC Flash
- White Paper: EMC XtremSF—Delivering Next Generation Performance for Oracle Databases
- White Paper: EMC XtremSF—Delivering Next Generation Performance for MySQL Databases
- White Paper: EMC XtremSF—Next Generation Performance for Microsoft Exchange 2010
- White Paper: Introduction to EMC XtremSW Cache