SSD PRIMER:
A BASIC GUIDE TO UNDERSTANDING SOLID-STATE DRIVES

Application Note

Document #AN0013 – Viking SSD Primer | Rev. A

Purpose of this Document

This application note was prepared to help OEM system designers evaluate the performance of Viking solid state drive solutions by using the same benchmarking methodology that Viking performs in it's SSD test facility. The SSD performance stated in the Viking SSD datasheets can be achieved by following the same Viking approach to SSD benchmarking which has been outlined in this document.

Table of Contents

Table of Tables

Table of Figures

1 Solid-State Drives Defined

A Solid-State Drive (SSD) sometimes referred to as an SSD flash drive, is made from semiconductor components known as “NAND flash”. NAND flash is rewriteable non-volatile semiconductor memory. There are several types of NAND flash and each type has a different price/performance measure associated with it. Flash drives are usually made with only one type of flash and are not mixed with other types of NAND.

All SSD’s include a flash controller chip to transfer data, manage flash operations and interface to the host computer.

SSD’s come in all shapes and sizes (a.k.a. Form Factors). The most common sizes are similar to hard disk drives used in servers, storage systems and computer laptops.

SSD’s are smaller, faster and more rugged than hard disk drives. Today, SSD’s are quickly replacing hard drives in applications where performance and environmental challenges matter most. They can outperform hard drives in many platforms, including laptops, desktops, storage systems and servers. The balance of capacity, high performance, and reliability make SSD’s the first choice for high speed data storage in enterprise, industrial and mobile computing applications as diverse as:

- Blade Servers, Rack servers, Storage arrays
- Tablet PC, Rugged/Mobile PCs/Laptops
- Network Cache Appliances, Telco Switches
- Event Recorder, Solid-State Recorders (SSR)
- Military/Aerospace Computers and Instrumentation
- UAV/UGV, Robotic Systems
- Single Board Computers (SBC), Vending machines
- Display systems, Video on Demand (VOD)
- Point of Sales (POS), Kiosk, Digital Signage

2 Flash Technology for SSD’s

A flash device in a single monolithic package is called a flash chip. Several flash chips are mounted onto a printed circuit board (PCB) inside the SSD to create large storage capacities.

2.1 Types of NAND Flash

There are at least 4 different types of NAND flash chips in use today; each categorized according to the amount of data a memory cell can hold and the length of its useful life.

- SLC - Single-Level Cell stores 1-bit of data
- MLC - Multi-Level Cell stores 2-bits of data

- eMLC - Enterprise Class MLC (High Endurance MLC)
  (Longer life versions of MLC that employ slower erase/programming time to lengthen life)
- TLC - Tri-Level Cell stores 3-bits of data
- QLC - Quad Level Cell stores 4-bits of data

MLC flash is the most common flash type used in SSD's, primarily because it has the lowest cost, often measured as $/Gigabyte. The other NAND types generally command higher prices due to their longer life, faster access times, higher reliability, and greater temperature tolerances. eMLC and SLC flash are usually found inside high-end SSD products used for industrial, military or business enterprise applications that require high performance under heavy usage (24 hours/day x 7 days/week).

2.2 Geometry of the NAND

The geometry of the NAND is the physical spacing between electrical circuits inside the NAND. As the spacing gets smaller, more memory can be packed into the chip making it less expensive. Today, MLC NAND is in the range of 20- 35 nanometers for memory capacities of 32Gbit and 64Gbit. However, one side effect of putting more memory into a smaller space is the reduction in the useful life of the flash. This is where SSD controller technology is critical and why SSD's from different vendors are not the same.

2.3 How NAND flash works

Flash data is stored as an electrical charge (voltage) in memory cells. Billions of memory cells are organized into "pages" of data. The pages are then accessed by the flash controller in "blocks" at the NAND interface. Flash is a random access medium in which the time to access a page of data does not depend on location (unlike a hard drive's sequential access nature).

2.3.1 NAND Interface

There are several choices for the NAND interface to the flash controller and flash chips are designed using only one type, which often depends on the manufacturer of the flash chip.

- Asynchronous (a.k.a. Legacy) access mode prior to year 2010
- ToggleDDR or ONFI access modes for year 2010+

2.3.2 NAND Attributes

While the technology of NAND has remained essentially the same over the last several years, the geometry continues to shrink while capacity continues to grow. The following table shows the migration of certain NAND attributes over the last 5 years.

Table 2-1: NAND Attributes

Attribute	2006	2007	2008	2009	2010
Max Die Density	2Gb	MLC: 8Gb	SLC: 8Gb/ MLC:16Gb	SLC:16Gb MLC: 32Gb	SLC:32Gb MLC: 64Gb
Geometry trend	9x	7x	5x	3x	2x
Geometry (nm)	90	72	50	34	25
Page Size (bytes)	2112	2112	4314	4320	8640
Block Size (KB)	128	MLC: 256	SLC: 256 MLC: 512	SLC: 512 MLC: 1024	SLC: 1024 MLC: 2048
Toggle Interface (MT/s)	SDR 40	SDR 50	SDR 50	DDR 166	DDR 200
Packages	TSOP (leads)	TSOP (leads)	TSOP (leads)	TSOP (leads), BGA (balls)	TSOP (leads), BGA (balls)
Temperature Range	Commercial 0 – 70c	Commercial 0 – 70c	Commercial 0 – 70c	Commercial 0 – 70C, Industrial -45 to +85c	Commercial 0 – 70C, Industrial -45 to +85c

2.4 Useful-Life and Shelf-Life of NAND

Flash is unlike other memory semiconductors because NAND flash has a finite useful life that wears-out over time as well as a limited “shelf-life” for holding data over long periods of inactivity. The term “flash endurance” is the useful life of the flash and the term “data retention” can be considered as the shelf-life of the flash.

There are several methods used to extend the longevity of SSD’s; the most important being the flash controller, not the flash, since endurance continues to drop as device capacities increase. The intelligence of the flash controller manages how often each flash memory cell is being used and maximizes the useful life of the SSD by “managing” the technical limitation of flash endurance and data retention. (See the section on SSD Operation)

3 Capacity of SSD’s

Capacity refers to the maximum amount of data an SSD can store and is described in 3 different ways.

1) raw capacity (unmanaged physical storage capacity of the flash)
2) usable capacity (raw capacity less any capacity reserved for the flash controller, over-provisioning etc.

3) formatted capacity (usable capacity less the capacity reserved for formatting and use with the operating system software)

Raw capacity is never completely available to the user, since flash must reserve a certain amount of capacity for the internal flash controller and its management tasks.

Usable capacity is the total amount of disk space available to the operating system software, but generally not to the user. The actual usable capacity is based on a level of “over-provisioning”, which is a portion of raw capacity reserved for the SSD flash controller to perform housekeeping and optimization of flash usage. (i.e. “wear leveling” tables to extend SSD life, spare flash cells to replace bad flash cells that accrue over time, file pointers and mapping, etc.) Today, SSD usable capacities are described in Gigabytes (GB) ranging from 25GB upwards to 2TB (Terabyte). It is typical for client grade SSD’s to be marketed with their raw capacity, however, there is always some amount of over-provisioning included. For example, a 256GB client SSD may actually have about 245GB usable space. Due to their usually lower performance and feature set, additional over-provisioning is not always useful. Client SSD’s marketed with their usable capacity are going to have a feature set closer to that of an Enterprise SSD. Enterprise class SSD’s are specified by their usable capacity. While an Enterprise SSD may contain 256GB of raw flash to ensure high performance and 5 years of endurance in datacenter applications, the usable capacity is more typically 200GB.

Formatted capacity, in most cases, is the true measure of capacity that is available to the user. It varies depending on the type of format and type of operating system. An example of how the actual formatted capacity can be found is by clicking on the disk drive icon (or ‘My Computer” in Microsoft Windows) to show its properties.

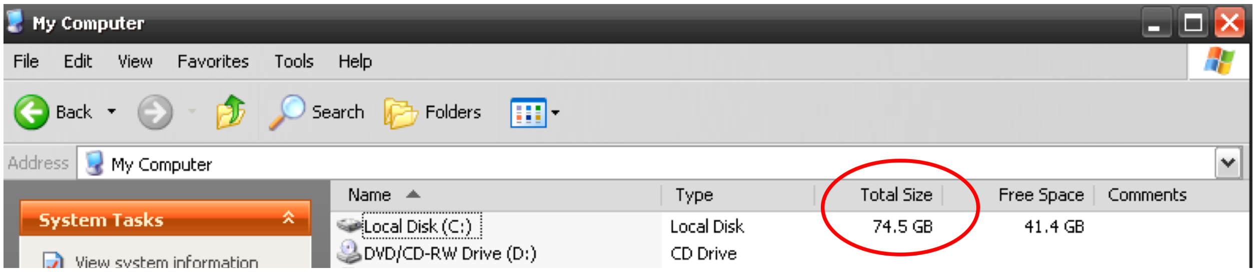


Figure 3-1: Example of Formatted Capacity reported by Microsoft Windows

4 Speed of SSD’s

The SSD data transfer rate depends on a number of factors including flash controller architecture, NAND interface, NAND type, and host system interface.

4.1 Definition of SSD Speed/Performance

Although "throughput" and "bandwidth" are common ways to describe performance, SSD speed is really based on two important read/write metrics:

1) Throughput between the host and SSD for sequential writes and reads to contiguous blocks of flash memory. Sequential bandwidth speed is measured in megabytes per second (MB/s).
2) Operational performance between the host and SSD for writes and reads to random blocks of flash memory. Random read/write performance is measured in the number of Input/Outputs Per Second (IOPS).

The most significant speed metric to an SSD user will depend on the intended application. Applications that involve frequent transfers of small and random amounts of data to the drive will benefit from high IOPS numbers which provide a quick response time to the host (i.e. virtualization, caching, databases, OLTP etc.). Applications that involve moving large sequential amounts of data (i.e. video streaming, data acquisition, data backup/restore, event recording, image processing etc.) will benefit from high throughput (i.e. MB/s).

The SSD flash controller can be setup using the operating system to optimize for IOPS or bandwidth (MB/s) in order to increase the speed of small data transfers or large data transfers, and to match the size and type of data access.

5 SSD Interfaces

There are several host system interfaces available for SSD's.

- USB - Universal Serial Bus
- CF - Compact Flash
- Cfast - Compact Flash with a Serial ATA interface
- SATA - Serial ATA
- SAS - Serial Attached SCSI
- FC - Fibre channel
- PCIe - PCI express
- Ethernet

SAS, FC, PCIe and Ethernet SSD's are usually for advanced users or business enterprises needing extra performance and reliability. Because of the sophistication of these drives, they are beyond the scope of the introductory document on SSD's. Contact Viking for more information.

5.1 USB Drives

The speed range of USB SSD's is dictated by the version of the flash controller used, as shown in the table below:

Table 5-1: USB SDD Speed Grades

USB	Gbps
Version 1.0	Up to 0.0015
Version 1.1	Up to 0.012
Version 2.0	Up to 0.48
Version 3.0	Up to 5

Notes:
Gbps, Gigabits per second

Later versions that are usually faster are backwards compatible to work in slower/older computer ports.

5.2 Compact Flash Drives

CompactFlash (CF) cards use an ATA / PC card / CompactFlash Interface and can support various data transfer modes. The transfer rates can be in the range of 40MB/sec or 3K IOPS. Compact flash cards measure slightly under 1.5 x1.75 inches and are available in two thickness.

- Type 1 (0.13 inches)
- Type 2 (0.26inches)

5.3 CFast Drives

The host system interface for CFast drives is SATA with additional enhancements to the specification such as sleep modes, current limiting and hot swap capability. CFAST drives use the compact flash card (CF) form factor with similar mechanical dimensions, but employ a different connector.

5.4 SATA Drives

Serial ATA (SATA) SSD drives support the ATA command protocol. Similar to USB drives, SATA drives can also operate at different speed grades based on the version of the flash controller, as shown in the table below:

Table 5-2: SATA SDD Speed Grades

SATA	Gbps
Version I	Up to 1.5
Version II	Up to 3.0
Version III	Up to 6.0

Later versions of SATA, have higher data transfer rates, but are backwards compatible to work at slower speeds in slower/older computer ports..

6 SSD’s in a System

The number of SSD’s that can be deployed in a system is only limited by the number of sockets (a.k.a ports) available. Most computers today come with multiple USB and multiple SATA ports making it possible the run several SSD’s at the same time, either internal or external to the system. However each SSD must be compatible to its port (SSD interface must match the host system interface for the port).

7 SSD Form Factors

Historically SSD’s, were designed with the same dimensions and packaging as the hard drives they were replacing. Today, however, SSD’s can break away from traditional form factors and take full advantage of the flexibility that Solid-State technology affords.

The popular “legacy” HDD form factors are:

- 1.8” (which is actually 2.1” wide)
- 2.5” (which is actually 2.7” wide)

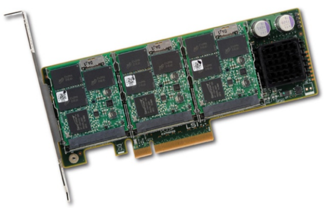
Newer form factors can shed the traditional plastic or metal cases and be incorporated into the system as bare board drives that are either reside in sockets or are soldered directly to the printed circuit board. Examples include:

- SSD’s mounted in the memory sockets of the computer (SATADIMM)
- Board mounted SSD’s hard soldered to the computer motherboard
- MO-297 (a.k.a. SlimSATA, a 2” x .5” bare board SSD with a socket)
- PCIe SSD (SSD mounted on the PCI Express computer bus in the computer
- MO-300 (mSATA) mini-PCIe SSD measures 50.80mm x 29.85mm x 4.85mm
- M.2

2.5”

1.8”

“MO-297 (SlimSATA)



PCIe card

mini PCIe MO-300 (mSATA)

embedded USB

Figure 7-1: Some Examples of SSD Form Factors

Note: Pictures are not actual size and are enlarged or reduced

Large capacity SSD are also available as separate external storage systems and can be connected to a computer or network of computers using a Fibre Channel or Ethernet interface. These “storage appliances” provide massive Solid-State storage capacities and use specialized architecture to provide redundancy and higher data reliability (i.e. RAID).

8 SSD Operation

The flash controller writes and reads NAND blocks over several pathways, called channels on the flash side and converts them into sectors that are accessible over a single pathway on the Serial ATA (SATA) host computer side as shown below.

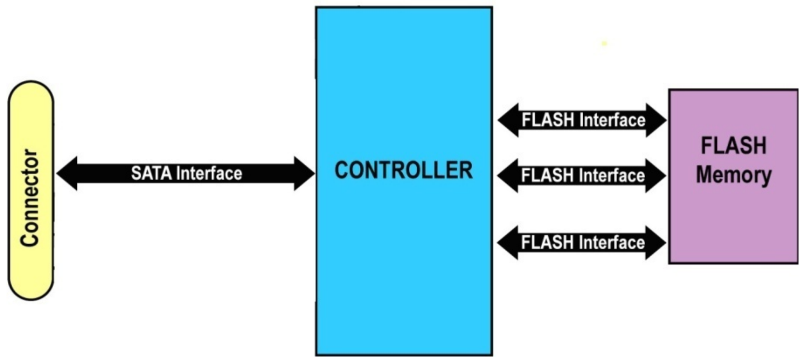


Figure 8-1: Simplified SATA SSD Block Diagram

The complexity of the flash controller to optimize the write/read performance, reliability and life of the SSD is a major differentiator among different SSD suppliers. Low-end SSD's often have low cost flash controllers that provide only basic functions for storing data under light or infrequent workloads and do little to increase flash life. These are typically referred to as client-grade SSD's.

Enterprise class SSD's perform under continuous operation, thereby requiring more sophisticated flash controllers and longer life flash that can withstand the higher-end workloads. The firmware and controllers for these SSD's are designed to ensure reliability under heavy usage , extended SSD life, maximum SSD speed (IOPS or throughput) and provide for advanced error detection and correction..

8.1 SSD Specific Terminology

To get a basic understanding of how an SSD works, it is important to understand its associated terminology. Many client SSD's do not have all these features.

8.1.1 Data Compression

Some controllers have the ability to compress data by analyzing repeatable data patterns and writing those patterns only once to the flash. Controllers with data compression capability can reduce the file size for data that is not 100% random (i.e. encrypted, zipped, PDF etc.) thus enhancing write amplification to minimize flash write cycles and increase SSD life.

8.1.2 Data Hardening

In the event of an unexpected power loss to the SSD, the flash controller and NAND memory will use internal self sustained power to finish any write cycles currently in process to the flash and also complete all internal housekeeping tasks like table updates. (i.e. this can occur on removal of the SSD from the computer or upon system power failure that kills power to the drive while it is actively writing data). Data hardening prevents data corruption that would result from an ungraceful shutdown of the flash, particularly MLC NAND, which is more susceptible to errors than SLC NAND.

Typically, a client grade SSD does not have this sort of data protection, however, enterprise class SSD's typically have the ability to self sustain power to finish these in-flight writes and gracefully shut down the SSD.

Flash memory devices exhibit unexpected behavior upon power failure. It is important to note that if a host write is in process when power is lost, that data will be lost and/or the contents of other MLC flash may become corrupted. Data that has been programmed or erased during a power failure can be unreliable, even if the data appears to be intact. Power failure can also corrupt existing data in flash prior to power failure. If possible, the host should issue a Standby Immediate ATA8 command (Opcode E0h) before powering down of the SSD

8.1.3 Data Retention (Shelf Life)

SSD's have a limited time (Shelf-Life) for holding data before it is automatically lost due to inactivity or after very long term storage. This is called Data Retention and it is the measure of time from when data is last written to the time when it can no longer be reliably retrieved.. Best case data retention is 10 years for storage of a drive that has had light or moderate use before reaching its end of life limit and the SSD storage temperatures is approx. 75F (25C). The worse case would be 3 months storage (inactivity) IF the drive has reached its end of life AND the SSD is subjected to storage temperatures of 100F+ (40C+) . Exposure to higher temperatures exacerbates the issue and shortens data retention. Generally speaking, SLC NAND flash will have a longer shelf life than MLC NAND flash.

8.1.4 Drive Life Protection (and Performance Throttling)

Some advanced flash controllers can be programmed to ensure that the life of the SSD will not expire before a certain period of time. This is accomplished by adjusting the speed of program/erase cycles (writes) to a level which will set the drive life to a predetermined amount of time; often the warranty period for the drive. Read operations are unaffected by this approach to protecting drive life.

8.1.5 Endurance (Useful Life)

Endurance is a measure of the maximum # of rewrites or erasures that the flash can withstand before it can only perform reads and can no longer perform writes.
When the endurance limit has been reached, the SSD is usually "retired" from service and can only be used as a read-only device for archiving or long term storage (i.e. for backup data). This is the point in time where data retention (explained above) is most critical.

Flash endurance depends on the type and capacity of the NAND flash inside the SSD, as shown in the table below.

Table 8-1: Endurance Trend

Characteristics	2006	2007	2008	2009	2010
Monolithic Die Density	2Gb	*MLC: 8Gb*	SLC: 8Gb MLC: *16Gb*	SLC: 16Gb MLC: *32Gb*	SLC: 32Gb MLC: *64Gb*
Geometry (nm)	9x	7x	5x	3x	2x
Endurance	SLC:100,000	*SLC: 100,000* *MLC: 10,000*	SLC: 100,000 *MLC: 10,000*	SLC: 100,000 *MLC: 5000*	SLC: 100,000 *MLC: 3000*
eMLC NAND (Enterprise class MLC) is ~ 10,000-30,000 write or erase cycles					

8.1.6 Garbage Collection

Garbage collection refers to a flash controller's housekeeping processes, transparent to the host computer, that recycle any flash space that contains obsolete data, which has already been modified and rewritten to a new flash location inside the SSD.

8.1.7 Native Controllers

A native flash controller is a single monolithic ASIC or FPGA, not needing a separate "interface" chip to communicate with the host computer. Native controllers can operate at the raw speed of the host without any interface "bridge" chip.

8.1.8 Over-Provisioning

Over-Provisioning is the difference between raw capacity and actual usable capacity based on a portion of the NAND that the SSD controller reserves for internal functions, such as storage of the controller firmware, flash housekeeping and usage optimization, flash endurance "wear leveling", garbage collection, spare block replacement of bad sectors, SSD index tables, file pointers, and logical to physical mapping. SSD over-provisioning also works like HDD short stroking (using only the outer/faster tracks on the disc platter). High levels of over-provisioning will lengthen the life of the drive at the expense of reduced usable capacity. Over-provisioning ranges can vary as a percentage of raw capacity.

8.1.9 Read Disturb

Excessive reads from the same flash memory cells can cause data loss from the nearby unread memory cells. Some flash controllers prevent Read Disturb data loss from multiple reads at the same flash memory address by tracking reads and performing read wear-leveling before cell structure is negatively impacted (i.e. repeated reads of the controller firmware stored in flash, repeated reads of system operating system software stored on flash etc..)

8.1.10 S.M.A.R.T

S.M.A.R.T. (Self-Monitoring, Analysis, and Reporting Technology) helps warn system software, the system administrator, or a user of an impending drive failure, while time remains to take preventive action.

8.1.11 Spare Blocks

Reserved drive capacity, known as spare blocks, are used to replace any sections of flash that have become defective.

8.1.12 SSD Index Tables

A section of flash that contains the logical addresses and their mapping to the physical flash addresses These index tables contain file pointers used by the operating system, to indirectly access the data using the SSD flash controller as an intermediary.

8.1.13 Storage Tiering/Pairing

Storage Tiering or Pairing refers to the practice of using SSD's and HDD's in the same system to gain the benefits of both types of storage devices (i.e. SSD's for fast boot up and high speed data I/O to frequently accessed data; HDDs for large, low cost storage/archiving).

8.1.14 TRIM

A host initiated software command to the SSD to allow the SSD to reclaim any flash space that has previously allocated to older files that have been deleted by the host.

8.1.15 Wear-Leveling

A flash controller housekeeping process, transparent to the host computer, that prevents premature endurance wear-out of any flash cell by evenly distributing Write/Erase cycles over the entire flash capacity. Wear leveling extends the endurance of the overall SSD by preventing repeated write/erase to the same flash locations; particularly relevant for file systems and file management algorithms that repeatedly write/erase identical logical locations which are physically mapped.

9 SSD's Features

There are many features of SSD's that are common to hard drives such as

- Data Encryption (AES-128, AES-256)
- Secure Erase
- CRC and ECC operation
- Bad block management
- Industry standard host interface

Since this document is a basic flash SSD Primer, any additional information on hard drive features that are available for SSD's, can be obtained by contacting Viking.

10 Frequently Asked SSD Questions (FAQ)

Q1: What are the benefits of SSD's over HDD's in the enterprise environment?
A1: SSD's can offer 10x-100x IOPS performance over HDD's and reduce the total cost of ownership (TOC) using fewer drives. This also reduces power consumption and cooling and the smaller form factor reduces space requirements.

Q2: What is the best usage for SSD's in the Enterprise storage environment?
A2: Storage tiering or pairing of SSD's with HDD's. Viking SSD's can cache high speed data I/O to frequently accessed data while HDDs provide large, low cost storage for archiving infrequently accessed data.

Q3: What is Viking's fit in the Enterprise SSD market?
A3: Viking can leverage it's expertise in creating small form factor modular solutions using controllers, DSP's, NAND flash and DRAM to create products having value and performance to fit our customer applications.

Q4: How will SSD's complement Viking's existing memory business?
A4: Viking will continue to market a variety of products to provide customer solutions. Please visit Viking's website at http://www.vikingtechnology.com for a complete overview of the Viking's product offering.

Q5: What type of SSD's are planned for the future?
A5: A Viking sale representative can provide you with our SSD roadmap for our SATA, SAS and PCIe product lines. These new SSD's are available in variety of form factors and performance levels.

Q6: What type of Flash memory is used in the Viking SSD's?
A6: Viking offers single-level cell (SLC) and multi-cell (MLC) NAND flash SSD's.

Q7: What is the reliability of the Viking SSD's?
A7: Enterprise SSD reliability is based on a number of factors including bit error rate (BER), Mean Time Between Failure (MTBF), write endurance and data retention. The values for these parameters are specified in the individual product datasheets.

Q8: How are Viking SSD's priced?
A8: Viking's SSD focus is on enterprise, industrial and military markets. We also offer lower cost client SSD's without power failure protection, less S.M.A.R.T. features and with lower random read/write IOP performance.

Q9: Who are typical Viking's SSD customers?
A9: Viking supports all Tier-1 and Tier-2, global enterprise OEMs and ODM as well as

leading industrial and military companies worldwide. Our mission is to develop and deliver high-technology products that optimize the value and performance of our customers' applications in the Networking & Communications, Enterprise Computing & Storage, Defense & Aerospace, and Embedded & Industrial markets.

11 Reference Documents

- Wikipedia information on SSD's
  http://en.wikipedia.org/wiki/Solid-state_drive

12 About Viking Technology

Viking Technology develops and delivers innovative high-technology products that optimize the value and performance of our customers' applications. Founded in 1989, Viking Technology has been providing Original Equipment Manufacturers (OEMs) with industry leading designs, engineering, product support and customer service for 20 years. For more information visit http://www.vikingtechnology.com.

13 Revision History

Related Products		Viking Element SSD Datasheets
Author		
Date		4-28-12
Document Number		AN0013
Date	**Edited by**	**Alteration to previous document revision**
9/14/17		Revise logo. Change company address

Global Locations

US Headquarters	India Office	Singapore Office
2950 Red Hill Ave. Costa Mesa, CA 92626 Main: +1 714 913 2200 Fax: +1 714 913 2202	A 3, Phase II, MEPZ-Special Economic Zone NH 45, Tambaram, Chennai-600045 India	No 2 Chai Chee Drive Singapore, 109840

For all of our global locations, visit our website under global locations. For sales information, email us at sales@vikingtechnology.com