New technique for data recovery from severe bad sectors
—— Super Image (ShadowDisk)

By SalvationDATA Technology, LLC.
16/08/2008

Summary: This article focuses on discussing the innovative technology of disaster data recovery - ShadowDisk technique, a very different technique from traditional image technology. By this technique, there is no need to create an image separately during the whole data recovery process. Instead, an index log will be created synchronously to record the sectors retrieved from a source disk on the counterpoint section of a ShadowDisk. Any repetitive read attempts onto the retrieved sectors will be turned to the ShadowDisk directly according to the index log. Thus, it can reduce read times on the source disk to preserve its original and intact situation. Furthermore, ShadowDisk technique can read any specific location of a source disk according to the value of the LBA, which traditional image technique is far to reach.

Keywords: Logical Block Address（LBA）, Data Compass（DC）, ShadowDisk, Super Image

1. Background

Along with fast development of informationization in 21st century, computers have been applied to commercially and individually aspects with a remarkable growing speed. So it makes various electronic information, namely, the data - - become important and valuable property which cannot get lost both for enterprises or individuals. How to guarantee the integrity, usability and security of storage data has then become a crucial question day by day.

Data loss can be summarized by two broad headings: Logical and physical causing. Accordingly, the two corresponding solutions are logical data recovery and physical data recovery. Logical data recovery usually is called as software-level file system restoration. In another word, it requires no mechanical or physical repairs, for example, problems like deletion or clone by mistake, virus, and partition loss and so on , they belongs to logical data recovery category. Its specificity can be manifested as these following symptoms: the disk cannot be read normally via the operating system; document could not be found or opened as usual; the partition is lost; or" unformatted partition" messages pop up and etc.

Correspondingly, physical data recovery involves mechanical or physical repair to retrieve data caused by damaged chips or components, such as scratches on platters, bad track, head crash, motor seizure, burnt PCBA and other components' malfunction and so on. Its specificity can be shown as: disk can not be detected via operating system, accompanied by sound like "tick tack"; or disk is recognized by system but cannot be read normally, or disk can not respond to system commands and freeze caused by the bad sectors.

Bad sectors can also be divided into logical bad sector and physical bad sector. The Logical bad sector is called as false bad sector; which is generally caused by softwares or users' improper operations. This

kind of problem generally can be solved by professional logical data recovery programs. The physical bad sector is usually factory defective ones. It indicates there are physical damages on track of disk. The only solution for physical bad sector is to reallocate partition or the sector.

To retrieve lost data caused by logic problems or bad sectors, it is very easy to resolve through general data recovery software or even through manually amending certain parameters of disk. However, it is more troublesome to recover data from disk caused by physical problems. Problems as damages on Head stack, spindle motor or chips on PCBA, can be normally solved and detected by operating system after replacing damaged components. Then the procedure moves to easy file system restoration. Nevertheless, for a disk with physical bad sectors, even if the operating system is able to identify it normally, but during the period of operating system initialization, the disk is likely to hang unresponsively or freeze; you can not read disk data through the operating system at all. So far, there is no one effective solution to this problem. Most data recovery companies try to solve the problem by extracting data and creating an image disk. Unfortunately, the abuse of image technology often brings some unexpected problems and serious consequences as follows:

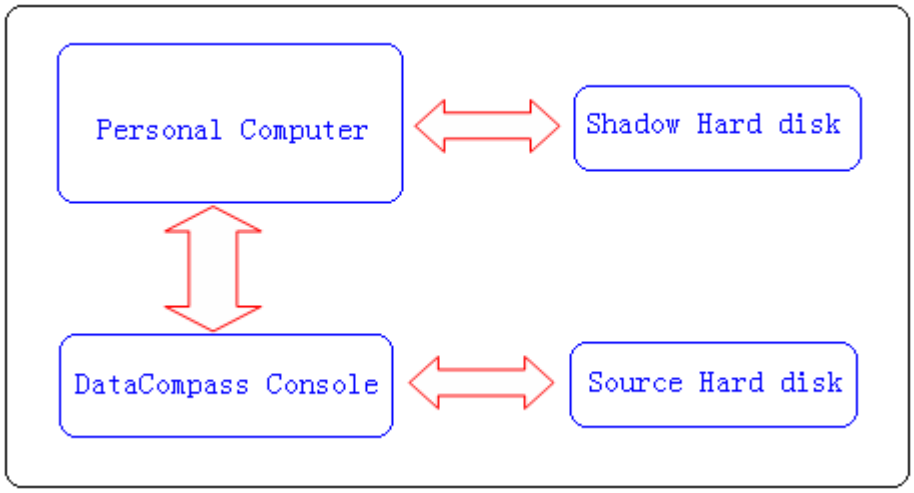
- Image technology can not effectively read the full spectrum of LBA, alternative between a certain partition and the entire disk.
- Image technology can not solve the problem by reading data from bad sectors effectively, long time read attempts to certain bad sectors by force, it can cause head injury as well as severe further damage on source disk.
- Image technology takes long time to copy data, sometimes it lasts for a couple of hours or even a full day. The long read time is more likely to increase the bad sector on source disk.
- Image technology consumes system resources excessively and you can not make use of other operations.
- Image technology may activate the trigger of Grown Defective List of drive (G-LIST) because it is trying to read bad sector repeatedly, which will result in firmware error, and the disk may not be identified correctly due to G-LIST overflow.

To this end, Salvation DATA technology Inc. has released a complete solution specifically designed to solve this series of above problems—Data compass. It adopts the unique creative technique: ShadowDisk, an innovation different from traditional disk image technology, which can solve these following problems: 1) the patient disk hangs or freezes while reading through operating system or it can not be read normally; 2) read data from physical damaged disk fast and recover data efficiently; 3) Recovery data from a logical structure corrupted disk; 4) make cumbersome, inconvenient and complicated operations much easier. All these advantages empower Data Compass to work out the following situations:

- Need to recover data from a physical damages disk
- Need to recover data from a logical structure corruption disk
- Need to recover data from a disk with both physical and logical problems

2. Solution

In order to effectively solve the unresponsive problem happened when the operating system directly read the defective disk source disk), and reduce the occupation of the system resources, this technology adapts a Data Compass to indirectly connect the source disk with the personal computer instead of a direct connection. This indirect connection solves the problem caused by the long times' no-responsive and hang caused when the every scan towards the source disk.
The whole process and the framework to solve this problem showed as below:



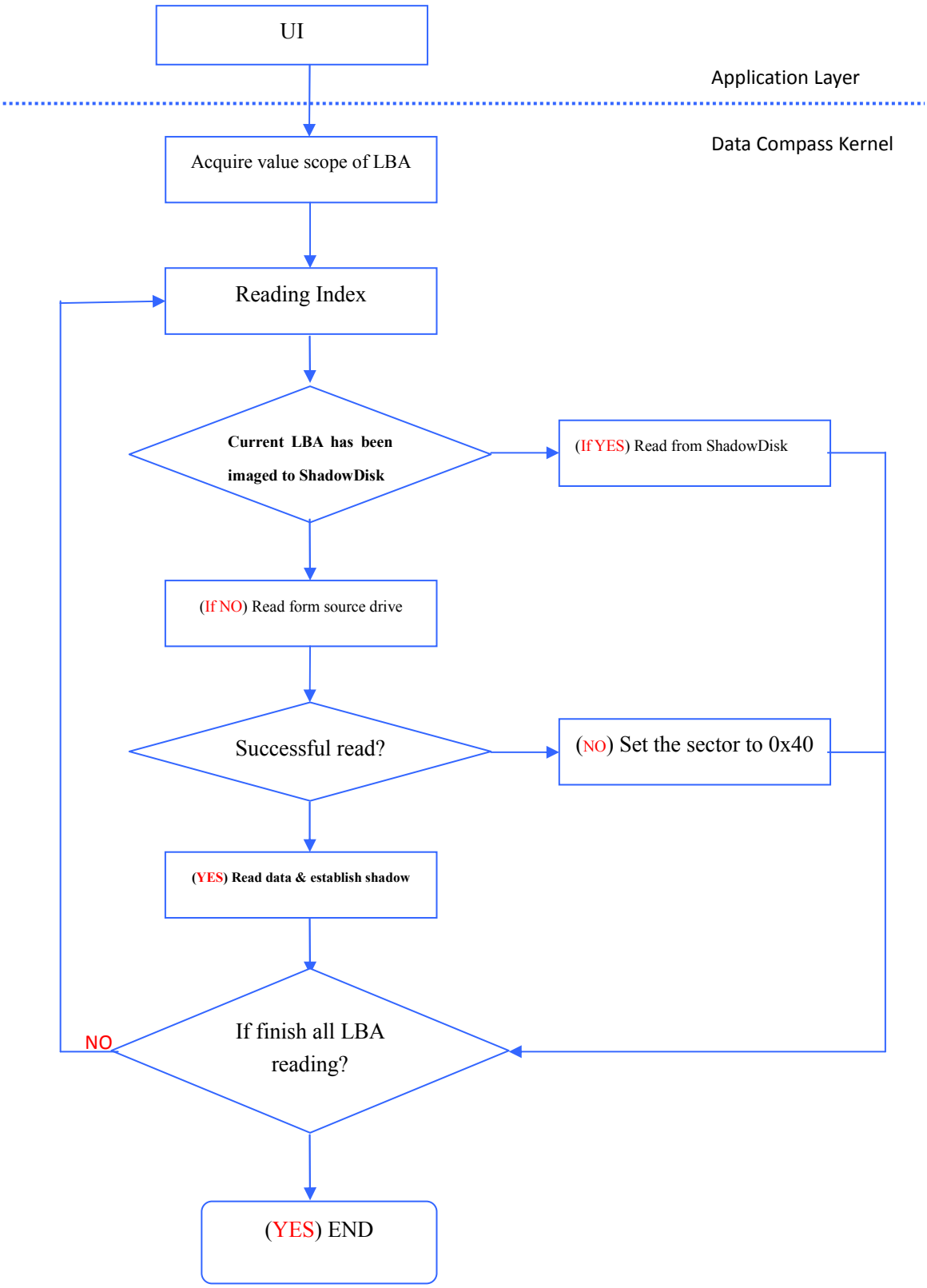
Whole Framework

From this framework, we can clearly note that the source disk is directly connect with the Data Compass, then through a USB cable the Data Compass is attached to the personal computer but not in the direct connection between the source disk with the PC. Obviously, every read and write command from the system must first arrive to the Data Compass. Then Data Compass will execute some corresponding operation based on these commands.

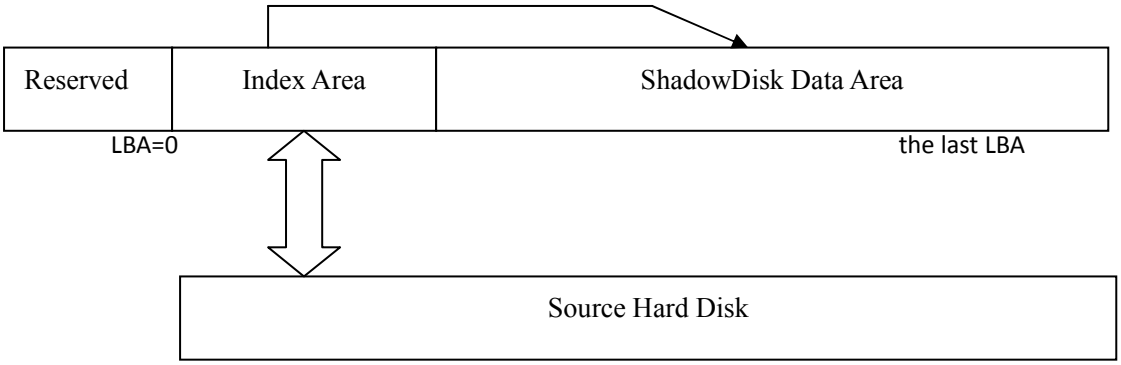
At the same time, you can see there is a shadow hard disk in this process. This shadow disk is straightly connected with the personal computer to take full advantage of the effective resources in the system.
In this process, the function of the shadow disk is not directly as the imaging disk of the source drive. It barely contains an immediate indexical imaging data result from the optimized algorithm and protocol by Data Compass. The requirement for the capacity of the shadow disk surpasses the traditional imaging technology which requires a bigger capacity than the source disk. In this process, by this optimized algorithm and protocol, the capacity of the shadow disk can smaller than this of source disk. Of course, if you have a chance to get a big capacity hard disk, we suggest you to use this bigger one.

3. Technique Principles

From the above solutions, we can easily comprehend that Data Compass Console is the core of the whole system, with all the operations related to it directly. The realization principles flow chart is shown below:



Once the whole system installed and the DC begin to read the source hard disk, the Data Compass Console will automatically inspect and acquire effective parameters from the source HD. Then the ShadowDisk will move to format stage (Attention: the data in ShadowDisk will be removed in format stage so you should make sure that data in ShadowDisk is of no value or it's a zero-filled HD.) This so called initializing ShadowDisk is just to do something with it, generate some space as index area, other space of shadow data area. Therefore, the corresponding relationship of configuration and the data in source hard disk are as follows:



The condition of each bit in index area reflects whether a certain LBA in the source hard disk has established a corresponding LBA shadow in the ShadowDisk data area. In other words, it's a matter of whether the LBA value successfully read and backup. After the format stage, the ShadowDisk will be divided as Reserved, Index Area and ShadowDisk Data Area Synchronously all the bytes in Index Area will be cleared.

After the ShadowDisk formatting and the ShadowDisk function starts working, the Application Layer sends out requests to read data from source hard disk, then Data Compass Console make corresponding special disposals according to it. First, Data Compass Console will estimate the LBA is 1, then it means that the data in this LBA has formed a Shadow in ShadowDisk data area and backup. At present, the read of this data only need to read the data in the corresponding LBA in the ShadowDisk data area, no need to read again from the source disk so as to reducing the read times to the source disk and effectively protect the source disk. But, if the current state is 0, then it means that the data in this LBA hasn't formed a Shadow in ShadowDisk data area and backed up. So you must read this data in this LBA again from the source disk.

There are two potential results when reading the source HDD: successful or failed. If the source disk has been read successfully, DC will write 1 to the bit of the corresponding LBA in the index model of the shadow disk, and the data in this LBA will be established as a shadow in the corresponding data in this LBA to the ShadowDisk. In this way, the future read of data in this LBA will be directly from the ShadowDisk instead of the source disk. In case of unsuccessful reading from the source HDD, DC system will allow users do specially corresponding operation on the reading flow. DC has set the retry times in its default executing flow. If the data can be successfully read during the defined time, DC will execute the successful reading flow. That means DC will modify the state of the bit in corresponding LBA of index area, and built a shadow in the corresponding location in the ShadowDisk data area. If the data is unsuccessfully read within the defined time, DC will no longer read it, but set the data of

the corresponding sectors of this LBA as 0x40, and then store into the buffer. This operation can cheat the system and avoid the system automatically read the data in next LBA until all the LBA set by user have been read thoroughly. If the users set it as intensive reading, DC will change the physical property of head stack, which means DC will increase the power supply of the head stack to read the data. This operation leads to high success rat, but easily causes damage to the heads, so few adoption of this operation is suggested. Certainly, it is very important that you should clearly know the location of the required data when you decide to use this reading method. Nevertheless, you can use this method to recovery data when other methods failed.

From the entire flow which is executed by the principle, we know that the ShadowDisk technology is different from the traditional disk image technologies. The function of ShadowDisk is not simply used as image disk, isn't extract the data of the source HDD to the ShadowDisk one-to-one. The traditional disk image technologies should only recovery data after the whole image be completely established. However, by using the shadow technology, which needn't wait the shadow be completely established. Building the shadow and doing data recovery are carried on simultaneously. The shadow technology does not extract data directly from the shadow disk, it extract the data directly from the source disk. When do data recovery, first it will check the LBA of the resource disk has been established as shadow or not. If the shadow has been established, it will read the data from shadow disk directly. If the shadow hasn't been established, it will read the data from source disk directly. Also it will decide to establish the shadow or not by the result of reading. From the traditional image technologies, data should be extracted after the whole image is successfully established. Also the traditional image technologies are too week to read the LBA data from the parts where include some bad sectors, and even affect the data extraction on the parts where don't include sectors. This problem will never excite by using the shadow technology. Towards the failed reading LBA, it will not establish the shadow, it will just do special handling—fill full of 0x40 to cheat the Window system. As that the Window system will not do auto scan, which can prevent the Window system halt from reading the bad sectors.

The other advantage of using shadow technology is: the operation is so convenient, so easy. Users can establish the shadow at the discretionary setting area of LBA. Users can close the program at any time, no matter the shadow has been established or not in a certain time. When users do the operation again at next time, it will not rebuild the shadow which had been built. It will read the data directly from the shadow disk. It will only read LBA which hadn't been established as shadow. And will establish shadow of corresponding LBA.

Conclusion:

For the data recovery on the Hard Disk with bad sectors, the functions and advantages by adoption of shadow technology are quite obviously. This technology not only protects the source disk very well, but also has the following merits: very effective data recovery, quite high execution rate, very easy to operate, extremely convenience, no need much interference of user, etc..

4. References

1、 http://www.it13.com
2、 http://www.ntfs.com