Solid State Drives: The Beginning of the End for Current Practice in Digital Forensic Recovery?

Graeme B. Bell, Richard Boddington, Journal of Digital Forensics, Security and Law 5, 2011.

http://www.jdfsl.org/subscriptions/JDFSL-V5N3-Bell.pdf

FAQ / Frequently Asked Questions (April 2011)

Hi everyone. Richard and I have had quite a few emails recently asking questions with common themes, so here is an FAQ guide with some more information about our SSD research. I hope you find your answer here.

Below, SSD means 'solid state drive' and 'GC' means Garbage Collector, a disk-based program that resets unused sectors of an SSD to improve/refresh their performance back to normal levels - and coincidentally, permanently purges previously deleted data.

Q1. "I heard there was a paper that showed SSDs keep storing information even when you try to delete it. You seem to be contradicting that. So which result is true?"

The paper you've heard of is probably the FAST-11 paper by Wei et al (http://www.usenix.org/events/fast11/tech/full_papers/Wei.pdf). I can't speak for those authors, but I can point out some similarities and differences between the papers.

Short answer:

(I believe) their paper found that SSD drives have a mind of their own: when you try to force them to purge some data, they may not actually do it. Our paper also found that SSD drives have a mind of their own: when you would expect them *not* to purge some data, they may purge it anyway!

In short: whether you are trying to purge the data, or trying to preserve the data, the disk may do the opposite to what you expect. Both results may have serious legal consequences.

Long answer:

In the FAST paper, they are asking the computer operating system (e.g. Windows, or MacOS) to tell the drive to delete data. The OS tells the drive, but the drive has a mind of its own (technically known as a flash translation layer) and consequently does whatever it thinks best in the interests of performance and drive longevity. So attempts to overwrite logical addresses do not correspond to actual overwrites of physical addresses. If you crack the drive open and take a look, you find the data isn't really deleted. There's also the matter of some spare space in the drive which is used by the drive, but hard to access directly, called 'over-provisioning' - you get this even in hard disks.

In our work, the drive itself - not Windows - has the intention of physically wiping data from the drive, in order to improve performance. You can call this a 'pre-emptive reset'. So when data disappears viewed at the logical layer, it's because the underlying physical layer cells were pre-emptively reset by the drive. The clever bit is the drive is pretending to be like Windows (reading the NTFS data from the drive) in order to figure out what can be purged. The reason we can be sure that the physical data is really deleted is because the entire purpose of this 'garbage collector' is to prepare the physical cells for their next use as a means to improve performance.

There would be no point at all in the manufacturer adding the feature, if it did not permanently purge the data. Furthermore the resulting performance gains from pre-emptive resets of physical cells associated with cell resets are extremely well accepted in the SSD drive performance/benchmarking community.

Q2. Doesn't this SSD characteristic help criminals?

The Sydney Morning Herald, The Age, and Scoop.co.nz ran a story headlined "High-tech criminals outsmarting the law". The article was quite balanced, and had some interesting quotes, but the headline only represents at best half the problem.

The problem we found in our research was that SSD behaviour does not correspond well with the expectations/precedent of courts and forensic investigation. Data that would normally be preserved could be permanently purged from the disk, even while the disk is in the hands of e.g. the police, a company's forensic investigations team, etc.

That data might be evidence of someone's guilt, but it might also be evidence of someone's innocence. For example, consider each of these situations:

A. If, in the investigation of a case, prosecution experts found that a drive was substantially zero'd out in a very unusual way, they might classically begin to assume that someone was 'covering up data' - it might be taken as an early indicator of guilt and they might begin building their hypotheses around it. Not good if you're innocent to begin with. The problem here is that now, such an event will be happening simply by the normal actions of these SSD disks.

B. If a drive contained deleted (but not purged) files that could be used by the defence to prove their client's innocence, e.g. provide an alibi - but the SSD purges them while it is in the hands of the police because they handle the SSD inappropriately - then that would seem to be a potentially serious problem for an innocent on trial.

Our paper is trying to spread the word that the existing forensics process is now out of date, and has to be fixed (or that at least, courts and businesses need to be aware of weaknesses that have developed and will continue to develop). The existing process has - through technological change - become unsuitable in a manner that could change outcomes for guilty or innocent people.

Some thoughts from Richard:

"One problem that is evident, is that an over-reliance on digital evidence, circumstantial by definition, without independent corroboration can lead to false convictions and conversely, the guilty escaping conviction. Increasingly, there is less corroboration available in most criminal cases and the prosecution has more often to rely on 100% digital evidence. Look to my previous paper on digital evidence validation. I hope to publish more on validation issues before the year end that look for corroboration of evidence to ensure the forensic examiners does the best for the courts." ...

"I am not so much concerned that the guilty go free because of a lack of recoverable evidence, but that the innocent are stigmatised by the same argument: the absence of evidence is not proof it did not exist. For example, erased evidence may be exculpatory and exonerate a suspect. There seems to be a reversal of natural justice where innocence must be challenged by the prosecution but now the offender must prove they are not guilty - semantics perhaps but not for those falsely accused. Take Wiki leaks for example." ...

"I have testified at a case where the evidence implicated others and not the defendant and I was able to provide digital evidence that corroborated witness alibi. The jury must have been impressed with the evidence and the defendant was found not guilty. Exculpatory, deleted evidence was recovered by me and evidently by the police yet they put the poor chap through an expensive and horrible ordeal. " ...

Also, on a practical level, there are tools easily available to the public that are vastly more effective than SSD accidental erasure, if your goal really was to deliberately purge the data as quickly as possible. No, we will not tell you what they are.

Q3. You ran this experiment on one disk multiple times - but why not ten disks / 1000 times / a wide range of situations / 100 computers ...

First of all, to demonstrate that a previously undiscussed ‘grey area’ exists, it was only necessary to show a single example of the forensics process going catastrophically wrong. We’re sounding the alarm, not setting out to catalogue the entire problem authoritatively. Indeed, given the pace of change in storage technology, it may not ever be possible to authoritatively map out the shape of this grey area. It will be continually evolving, as drives get smarter and become more like independent mini-computers. Garbage collection, TRIM, data de-duplication, full-disk encryption, an increasing disconnect between the logical sectors and physical chips, filesystems that aggregate over multiple disks, cloud storage ... it’s going to be a mess as these become increasingly commonplace.

Secondly, while we want to see the experiment run on hundreds of disks, in hundreds of configurations, this is best achieved by having hundreds of other people involved and running the experiment for themselves, so as to avoid any systematic error that we might introduce ourselves. Science is best verified by communities, not teams or individuals.

Thirdly, there is a dangerous trend in modern science which takes experiments out of the hands of normal people, and restricts it to a small number of well-funded groups. Often this is the result of an experiment so grand that no one else can afford the time or money needed to reproduce it. I wrote a paper raising the alarm about this problem last year (http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2739/2456).

Grand experiments can seem impressive, but consider that they are not likely to be double-checked. Worse still, anyone who does reproduce it on a smaller scale and discovers disconfirmatory results (e.g. they can’t reproduce the original finding) risks being discredited merely because they did not match the scale of the original experiment. This is not a good situation for science. The first result should *not* try to be the most authoritative, it should instead try to be the most easily reproducible, to get the ball rolling. The most effective way to achieve this is to make the first result small in scale and cheap/easy to reproduce. Scale and breadth of findings are best achieved by a community, not a single team.

Fourthly, a grand experiment might have drowned the reader in technical or statistical detail, and obfuscate our point rather than clarify. We hoped that by writing the paper in the way we did, this idea, this experiment, and these consequences will be comprehensible to everyday people, such as members of jury or business managers.

Q4. Why not just have the police/forensics officer rip open the drive, read the physical chips, re-solder a new controller, write new firmware onto the drive without GC, ... ?

The problems here are law and education, rather than technical issues. We would agree, for example, that swapping out the drive controller for a near-identical one which does not have GC, but which is capable of understanding the flash translation layer, would avoid the risk of the GC running during the forensics process. It would be technically challenging and fraught with the potential for catastrophic error if there is only one physical drive to work on, but it could certainly be done if substantial funds were available to support it.

The problem is that currently there is no legal precedent or agreed process for such an act. Even if the prosecution/defence were to haul out the ‘brains’ of the drive and replace it with something of their own, that could look a lot like ‘tampering with evidence’, or even ‘embedding’ evidence that would not have existed in the normal operation of the equipment. These are legal issues that will take time and thought; it’s not as simple as just soldering/desoldering/reverse engineering the problem away.

There's also the matter of training. It could be said that students of digital forensics have had it easy for the last few decades - simply hook up the drive to a USB write blocker, run the imaging program, and you're done with the basics - the physical steps can be taught quite easily in class in an hour. But when you start talking about identifying firmware accurately, desoldering chips, writing new firmware to a drive, you're talking about actions where a single mis-step or error destroys the evidence totally and which require more specialist skills to carry out. That is very far from practical.

Q5. How do you know the data was physically deleted in your experiment? What if it only disappeared at the logical level? Why not strip out the chips and check?

a. *It seems very odd to suggest that GCs don't GC.* The whole point of the garbage collector is to reset cells pre-emptively to avoid delays later when data is being written to the drive. It would be quite insane for a company to go to the expense and risk of adding a complex and risky data-modifying garbage collector that does not actually collect garbage, but instead, simply pretends to. Recall that the GC is running on the disk with a simultaneous view of the logical filesystem, the logical sectors, and the physical chips, unlike an OS, and the GC is tasked to make physical sectors empty of data so they are ready for new data.

b. *Reproducibility.* The way we've conducted the experiment, anyone can try it at home or in their office and check our result. You don't need a big budget or any special equipment or electronics knowledge - we've supplied the configuration and software at the end of the paper.

Generally, experiments that require unusual equipment or technical skills, don't tend to be repeated. That means we are less likely to have people try to reproduce the work, and it makes it harder to persuade people of the result. Nothing persuades people more effectively than an experiment you can watch taking place right in front of you, using simple point/click actions and a little typing.

c. ***It's a non-issue, anyway.*** In real life, forensic investigators don't extract the chips and read them directly, except where they must - e.g. where someone damaged the disk by submerging it in water, smashing it to pieces etc. So, the question of physical deletion is a non-issue given present practice, where 99% of cases will involve simply reading logical sectors or logical filesystem from the disk with a write-blocker and the drive's own controller/firmware.

Q6. Are you sure it was GC and not the motherboard / TRIM / accidental writes?

Believe me, when you watch (in real-time) a disk with a professional forensic USB write blocker attached do nothing for a few minutes, then suddenly start purging data from the logical sectors at a rate of hundreds of MB/second in front of your eyes, it's akin to watching a robot car drive itself. There's no question that something rather spooky and out-of-the-ordinary is going on.

We've described the setup in full and so you can check for confounding factors yourself directly. TRIM, SSD aware motherboards, and accidental writes can be ruled out immediately based on the configuration we used.

Q7. "I could have told you that would happen" / "This has been possible for decades"

Agreed, it was obvious. We're well aware of this and we undertook the paper with clear expectations about what we thought we would find - and those expectations were broadly confirmed.

But talk is cheap, and ultimately not worth much in a court-room or science lab.

To actually experiment for real, and find out for sure what would happen: purging or not, and to what extent, and when, and how quickly, in the presence of a write-blocker - to document that thoroughly and carefully in a manner that can be easily reproduced - to establish a realistic yet simple scenario (a quick format is an activity that could be carried out in seconds, by either a naturally innocent or guilty person) - and to simulate a forensic extraction using a court-approved forensics expert and court-approved forensics equipment... that is more difficult than just ‘talk’.

Q8. How does this relate to TRIM?

TRIM is another issue which will complicate matters in that it will again not be possible to attribute purging of data permanently to an intentional effort to hide evidence. TRIM commands are special new disk commands that are sent from some operating systems or application programs to drives, which indicate that the operating system no longer needs the contents of those sectors.

Acting on that information by purging the un-needed data is *optional*, from the drive’s point of view, according to the ATA TRIM standard.

This means that again we have a problem where the drive can go off with a mind of its own - perhaps when in the hands of a forensic investigator - and purge all or part of the previously deleted data on the drive.

TRIM is well worth learning about, but keep in mind that it’s quite distinct to pro-active GC for non-TRIM operating systems. TRIM is going to be very important in future operating systems - it’s supported already in Windows 7, and should be available in the new Macos Lion (10.7). This is another major reason why we feel it’s the beginning of the end for forensic recovery of deleted data.

Q9. What happens if you delete files in your experiment, instead of quick format?

Well, performance enthusiasts on storage web forums can tell you that the performance characteristics of SSD drives with GC change slowly over time, often changing after 30 minute idle periods. However, such behaviour does not always seem deterministic or reproducible - in other words, it’s hard to show it happening in a way that suits a journal paper.

We did not formally experiment on that situation, since our goal was to establish a simple, realistic, reliable, easily reproducible scenario.

Also, informal discussions on this topic generally use indirect evidence such as the write performance of the disk to establish whether the GC has activated. Here, we prepared a ‘disk sampling’ program that continually samples direct measurements of the data content of the drive in near real-time and makes the GC’s behaviour directly observable.

We encourage you to try an experiment for yourself, and find out what happens.

Q10. “What disk did you use?” / “Did you check if TRIM was running” / etc.

Please read the paper in full, including the appendices, if you’re curious about this sort of thing, it is all described there. If you think we might have missed some variable that you believe affects the experiment, there’s no point in getting into an discussion about it without real data. It would be much better if you could set up this experiment at home using something close to the equipment that we used, and re-run it with the parameter you’re thinking of, and send us the details of any substantially different result you find (or publish the result yourself, if you prefer). We look forward to seeing your own results!