Chapter 26

TIME ANALYSIS OF HARD DRIVE IMAGING TOOLS

Jack Riley, David Dampier and Rayford Vaughn

Abstract Computer hard drives often contain evidence that is vital to digital forensic investigations. However, an authenticated working copy or "forensic image" of a suspect hard drive must be created before any data can be analyzed. As the capacities of modern hard drives increase, the time taken to create a forensic image, let alone analyze the data, increases significantly. This paper investigates two popular hard drive imaging tools, ICS ImageMASSter SOLO III and Logicube Talon. The results of the imaging experiments and timing analysis provide valuable guidance on selecting the appropriate imaging tool for digital forensic investigations.

Keywords: Imaging tools, hard drives, time analysis

1. Introduction

Digital forensic activities, at the highest level of abstraction, can be grouped into three basic tasks: acquisition, authentication and analysis. Acquisition involves seizing media and equipment that might contain digital evidence and processing the items to recover the evidence. During this process, at least two copies of all source media are made for purposes of analysis; the original evidentiary items are then catalogued and stored securely. Authentication is necessary to prove that the working copy of the digital evidence used for analysis is identical to the original. This is generally done by computing cryptographic hash values of the original and copy; the integrity of the copy is verified when its hash value matches that of the original. The final process, analysis, explores the copies of the original media to identify potential evidence and provide corroborating support for non-digital evidence. This pa-

per focuses primarily on acquisition and secondarily on the process of authentication.

Hard disk imaging devices create exact (bit-for-bit) duplicates of an original hard drive and, at the same time, calculate a cryptographic hash value of the original and copy. The time requirements for imaging hard drives is a serious issue, especially as cases frequently involve massive volumes of digital evidence and hard drive capacities are increasing significantly. Meanwhile, new demands on evidence acquisition are imposed by legislation such as the Sarbanes-Oxley Act, which requires mandatory document retention [3]. Since the analysis of digital forensic data is time intensive, time saved during the evidence acquisition phase can be leveraged during the analysis phase.

Several hardware and software tools have been designed for imaging hard drives, but they have greatly varying capabilities. Few, if any, researchers have analyzed the time requirements for these tools using rigorous experimental methods. This paper investigates two of the most commonly used hardware-based imaging tools, ICS ImageMASSter SOLO III and Logicube Talon. In particular, it describes the results of imaging experiments and timing analysis. The comparative study provides valuable insights into the performance of hard drive imaging tools and offers guidance for tool selection in digital forensic investigations.

2. Background

Hard disk storage capacities have increased significantly over the past ten years. Currently, two terabytes of storage can be purchased for under $1,000 [4], and the cost per terabyte of storage continues to drop rapidly. The availability of massive volumes of inexpensive storage is a boon to all types of computer users, but it also serves to increase the amount of electronic evidence that the digital forensic investigator has to sort through in civil and criminal cases.

Digital forensic investigators need efficient methods for acquiring and analyzing data. Roussev and Richard [4] report that one of the most widely accepted forensic examination systems took more than four days to organize case data on an 80 GB hard drive. Their results indicate that this was mainly due to I/O limitations of large capacity drives. Only after a case is opened and the data is indexed can investigation and analysis proceed. These steps also take a considerable amount of time, especially for cases involving data in the order of terabytes.

Several researchers have focused on making forensic analysis more efficient. Dandass [2] has used field programmable gate arrays (FPGAs) to implement pipelined pattern matching algorithms for speeding up the

search for image files on hard drives at line speed. In laboratory tests, the FPGA implementation required a little over 600 seconds to locate 24 image files placed in random clusters on a 40 GB hard drive; in contrast, state-of-the-art software-based methods running on a Pentium 4 2.8 GHz computer under Windows XP took more than 4,700 seconds [2]. The FPGA implementation also supports sector-by-sector copying of hard drives at speeds approaching 6 GB per minute.

Roussev and Richard [4] have sought to reduce the time needed for forensic analysis, especially in the face of the slow linear growth of I/O systems compared with the exponential growth of CPU performance and data storage capacity. Their research has shown that it is more efficient to access a hard drive once and perform analysis from a cached copy in memory. Unfortunately, the standard practice of using a single forensic workstation does not allow for a cached copy of any significant size to be analyzed due to constraints on the memory capacity of a single system. Roussev and Richard obtained good results using a specialized, distributed approach to forensic analysis. In particular, their approach produced significant reductions in data preprocessing and search times.

Unfortunately, these research results, while promising, have not yet transitioned to forensic practice. A need still exists for a practical technique to efficiently copy vast amounts of data in the least amount of time. Our work evaluates the time requirements of two leading hardware drive imagers, with the goal of assisting practitioners in choosing the right tool for an imaging task.

3. Experimental Design

Our experiments on hard drive imaging tools were designed to evaluate the base times required to create exact authenticated copies of hard drives. Two imaging tools, the ICS ImageMASSter Solo Forensics III and the Logicube Talon, were used in this study. Non-imaging functions provided by the tools, including hash value checks, were disabled or disregarded when the timing data was collected. Since timing display capabilities were an unknown variable in the study, a software stopwatch [1] was used for timing purposes. Using a stopwatch introduces human reaction time error, however, the error was assumed to be consistent and was minimized to the extent possible. In any case, the time measurements made in the experiments were much larger than the fractions of a second introduced by human error.

The timing analysis was conducted using ImageMASSter and Talon for one-to-one drive transfers. The experiments were carried out in two stages: (i) one-to-one IDE trials, and (ii) one-to-one SATA trials. The

IDE and SATA trials were both conducted using 80, 120 and 250 GB drives. These sizes were chosen because they are representative of the drives used in digital forensic investigations. Each trial involved ten iterations per drive, and both the tools were tested on the same drives.

3.1 IDE Hard Drive Trials

This section presents the results obtained using the Talon and ImageMASSter tools on 80, 120 and 250 GB IDE hard drives.

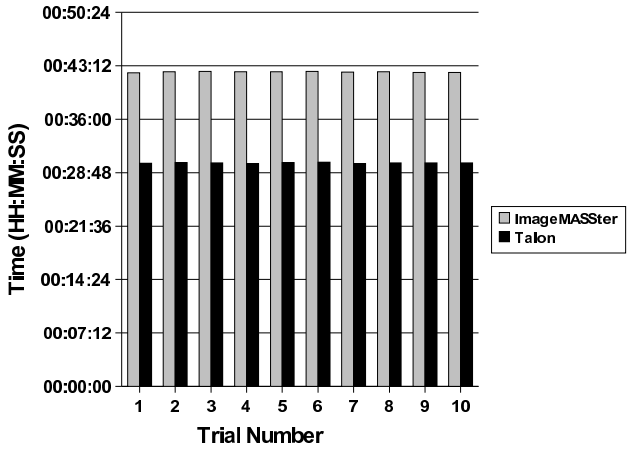

Figure 1. 80 GB IDE trial results.

Figure 1 presents the results obtained for the 80 GB IDE hard drives. Note that the experiment used two Western Digital WD800BB hard drives as the source and destination drives. The data shows that Talon outperforms ImageMASSter by about 30%, with an average time over ten iterations of 30 minutes, 7 seconds as opposed to 42 minutes, 22 seconds.

Figure 2 shows the results for the 120 GB IDE drives; 120 GB Seagate Barracuda 7200.9 drives were used as the source and destination drives. Once again, Talon was faster than ImageMASSter, with an average time over ten iterations of 45 minutes, 24 seconds compared with 54 minutes, 32 seconds. These results show a difference of 9 minutes, 8 seconds, which is less than the difference achieved for the 80 GB IDE trials; this indicates a possible scaling factor.

Figure 3 presents the results for the 250 GB IDE pair of hard drives (the source drive was a Seagate Barracuda 7200.9 and the destination drive was a Western Digital WD2500). As before, Talon is faster than ImageMASSter; however, the difference in performance is significantly less than that observed in the 120 GB IDE trials. The average times

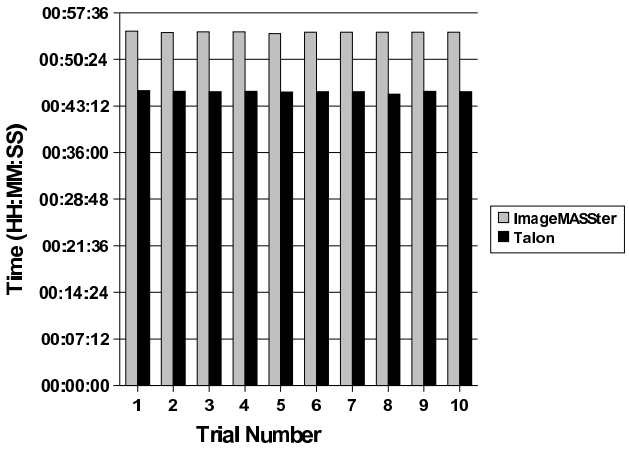

Figure 2. 120 GB IDE trial results.

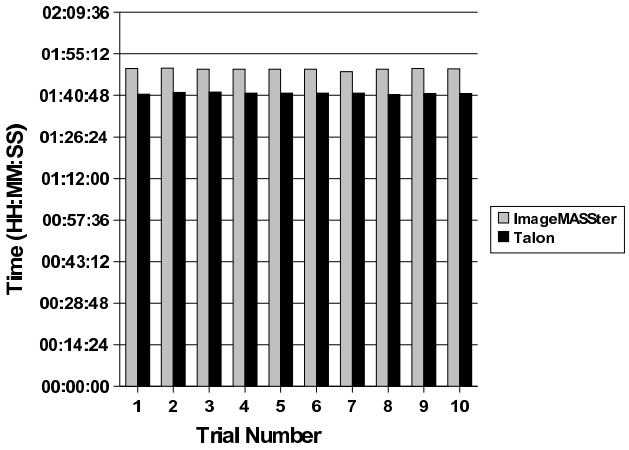

Figure 3. 250 GB IDE trial results.

over ten interations for Talon and ImageMASSter were 101 minutes, 32 seconds and 110 minutes, 8 seconds, respectively.

3.2 SATA Hard Drive Trials

This section presents the results obtained using the Talon and ImageMASSter tools on 80, 120 and 250 GB Serial ATA hard drives.

Figure 4 shows the results obtained for a pair of 80 GB Western Digital WD800JD SATA drives. Talon proved to be faster than ImageMASSter, with an average time over ten iterations of 31 minutes 18 seconds as opposed to 34 minutes, 39 seconds.

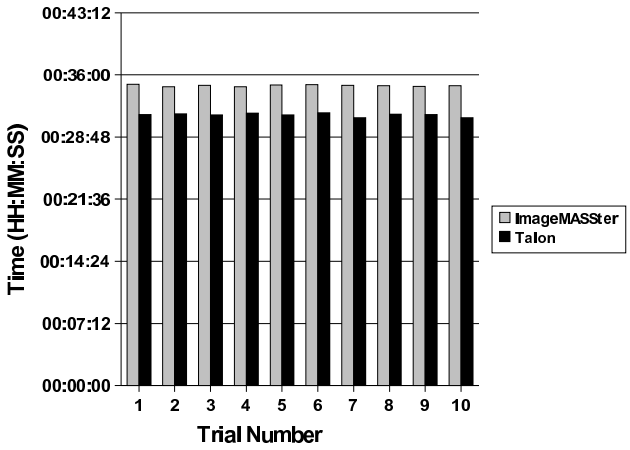

Figure 4. 80 GB SATA trial results.

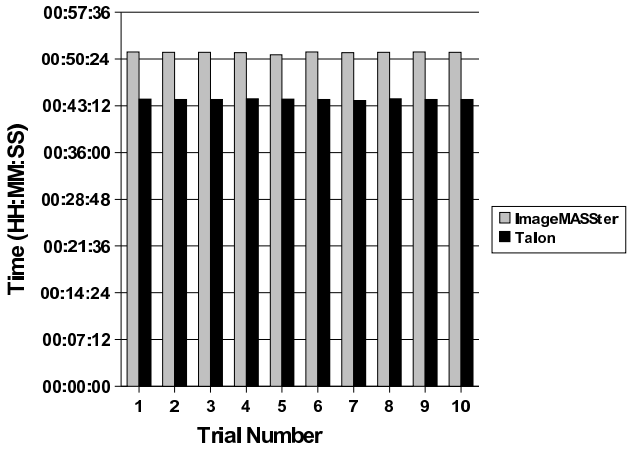

Figure 5. 120 GB SATA trial results.

Figure 5 presents the results for a pair of 120 GB Western Digital WD1200JS SATA drives. The difference between the average times over ten iterations in this experiment was 7 minutes, 13 seconds, once again in Talon's favor. The actual recorded average times were 44 minutes, 11 seconds for Talon and 51 minutes, 24 seconds for ImageMASSter.

Figure 6 shows the results obtained for a pair of 250 GB Western Digital WD2500KS SATA drives. The difference in average times over ten iterations between the two imaging tools was nearly 13 minutes, with Talon averaging 93 minutes, 45 seconds and ImageMASSter averaging 106 minutes, 39 seconds.

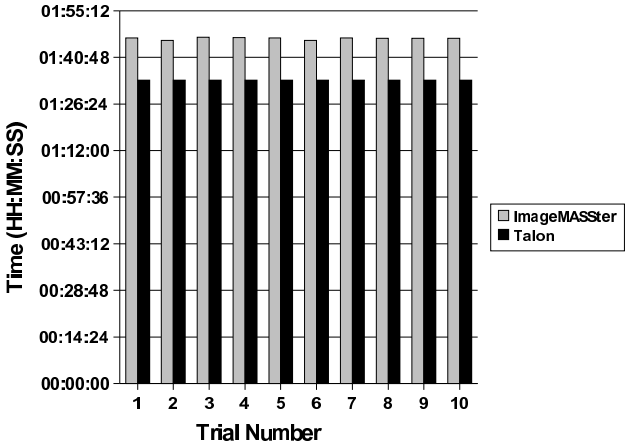

Figure 6. 250 GB SATA trial results.

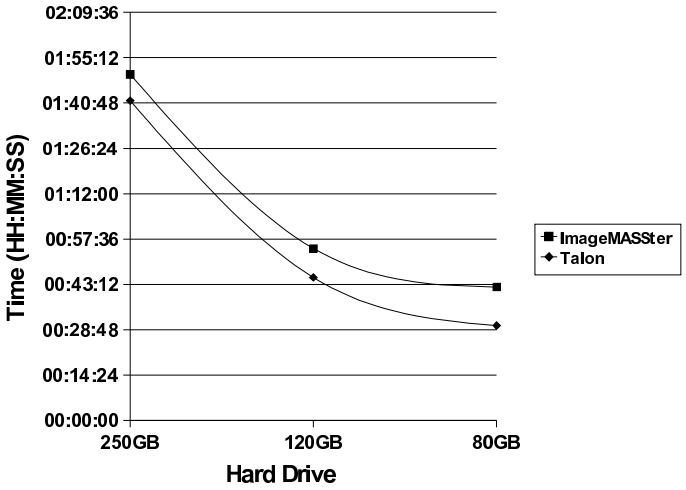

Figure 7. Average IDE drive imaging times.

4. Analysis of Results

The analysis of experimental data reveals three important observations. First, Talon performs better than ImageMASSter on IDE and SATA drives. Figures 7 and 8 show the average times for Talon and ImageMASSter; it is clear that Talon performed better on the drives used in the experiments.

Second, as seen in Figure 7, as IDE drive capacity increases, the difference in the imaging times for Talon and ImageMASSter decreases. The potential exists that, for very large drives, ImageMASSter may exhibit

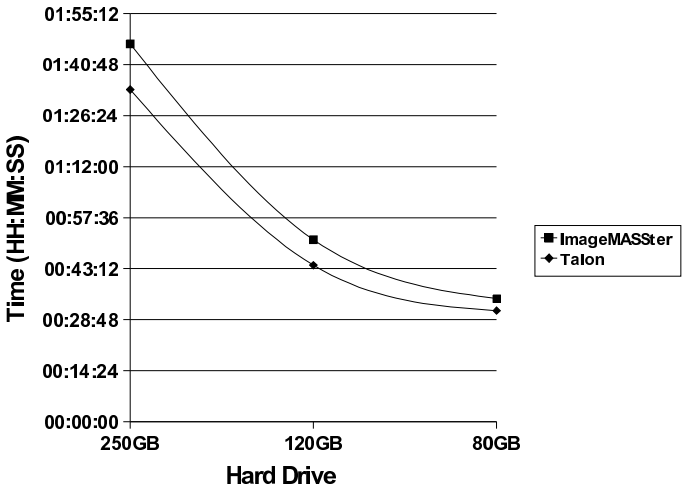

Figure 8. Average SATA drive imaging times.

better performance than Talon. However, Figure 8 reveals the opposite trend for SATA drives, i.e., the performance difference between Talon and ImageMASSter increases for larger hard drives.

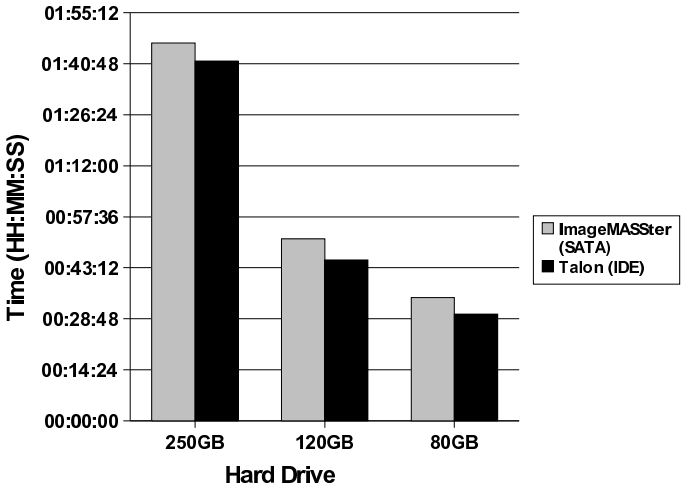

Figure 9. Average times for ImageMASSter (SATA) versus Talon (IDE).

The third observation is that not only does Talon exhibit better performance than ImageMASSter for IDE and SATA drives, but Talon actually performs better on IDE drives than ImageMASSter performs on SATA drives. This result, shown in Figure 9, is unexpected because SATA drives are supposed to be much faster than IDE drives.

Table 1. Computed t-values for test groups.

Test Group	Computed t-value
250 GB IDE	64.28
120 GB IDE	124.30
80 GB IDE	410.29
250 GB SATA	124.84
120 GB SATA	153.49
80 GB SATA	113.69

5. Statistical Analysis

A t-test was conducted to determine whether or not the difference in the results observed is statistically significant. The null hypothesis was that there is no significant difference between the data from the ImageMASSter and Talon trials. The respective means, standard deviations and variances were calculated for each data group. The variance for both groups of data was less than 6%.

Ten iterations were performed for each group (n1 = n2 = 10), yielding a degrees of freedom value of 18 (= n1 + n2 – 2) with p = 0.001. The proposed t-value for each group of data, referenced from the aforementioned degrees of freedom and p-value, was 3.92. The t-value was computed for each paired group of data using Microsoft Excel 2003; the results are reported in Table 1. If the computed t-value is larger than the proposed t-value, it can be concluded that there is a 99.9% probability of that the two groups of data are statistically different. Table 1 shows that the computed t-value for each test group exceeds the proposed t-value for the given p-value (0.001) and degrees of freedom. From these results, the null hypothesis can be rejected and the observed values are, in fact, statistically different.

6. Conclusions

Imaging speed is important to digital forensic investigators because of the large volumes of electronic evidence that are involved in civil and criminal cases. With storage capacities certain to increase in the future, the ability of imaging tools to quickly make authentic copies of hard drives will become even more critical. The experiments demonstrate a marked difference in the speeds of two popular hardware-based imaging tools, with the Logicube Talon outperforming the ICS ImageMASSter. While the difference in speeds might appear small, it is important to note that the time savings achieved when using a faster imaging tool can be significant for large capacity hard drives.

References

[1] Arantius.com, Stopwatch (tools.arantius.com/stopwatch).

[2] Y. Dandass, Hardware-assisted scanning for signature patterns in image file fragments, *Proceedings of the Fortieth Hawaii International Conference on System Sciences*, p. 268, 2007.

[3] M. Lange, Sarbanes-Oxley has major impact on electronic evidence, *The National Law Journal* (www.law.com/jsp/article.jsp?id =1039054510969), January 2, 2003.

[4] V. Roussev and G. Richard III, Breaking the performance wall: The case for distributed digital forensics, *Proceedings of the Fourth Digital Forensics Research Workshop*, 2004.