Digital Investigation 11 (2014) S36–S44

Contents lists available at ScienceDirect

Digital Investigation

journal homepage: www.elsevier.com/locate/diin

Information assurance in a distributed forensic cluster

Nick Pringle[*], Mikhaila Burgess

University of South Wales (formerly University of Glamorgan), Treforest CF37 1DL, UK

Keywords:
Digital forensics
Distributed processing
Media analysis
FUSE file-systems
Information assurance

ABSTRACT

When digital forensics started in the mid-1980s most of the software used for analysis came from writing and debugging software. Amongst these tools was the UNIX utility 'dd' which was used to create an image of an entire storage device. In the next decade the practice of creating and using 'an image' became established as a fundamental base of what we call 'sound forensic practice'. By virtue of its structure, every file within the media was an integrated part of the image and so we were assured that it was wholesome representation of the digital crime scene. In an age of terabyte media 'the image' is becoming increasingly cumbersome to process, simply because of its size. One solution to this lies in the use of distributed systems. However, the data assurance inherent in a single media image file is lost when data is stored in separate files distributed across a system. In this paper we assess current assurance practices and provide some solutions to the need to have assurance within a distributed system.

Introduction

The notion of using distributed processing to address the increasing scale of forensic investigations was first considered in "Breaking the performance Wall" in 2004 (Roussev and Richard, 2004). Despite being revisited several times since, (Ayers, 2009; Beebe, 2009; Garfinkel, 2010; Pringle and Sutherland, 2008; Richard and Roussev, 2006; Richard et al., 2007), this has not been developed and adopted as a workable solution. There has been a resistance to the idea of using an architecture where the data is moved and stored on a multitude of hosts for processing. In this paper we briefly consider the technical issues but conclude that the most important reason is the lack of a forensically sound approach to ensuring information assurance within a distributed system. This is required to ensure evidence management is regulated and clearly accountable for the legal community.

We will introduce our design for a middleware distributed processing solution, FCluster, which is specifically designed to provide assurance for the integrity of data.

Background

As digital forensic investigation methodologies have matured to accommodate the developments in technology, crime and investigative capabilities over the last 20 years, internal controls have been introduced to provide assurance standards required by the legal process.

Within our expectations of assurance there are a relatively small set of acceptable and 'trusted' investigative tools. FTK and EnCase are two of the most popular and trusted tools for digital media forensics. We know from more than a decade of use that their design endows confidence in the investigative process, and this is supported by these tools being tested for forensic appropriateness by NIST. In particular, the risk of 'mixing up data' between the evidence media and the host computer is negligible. There is no realistic way that data from another image could be introduced because there is no mechanism, other than operator error working on the wrong image, for this to

[*] Corresponding author. Tel.: +44 0 1443 4 83261.
E-mail address: nicholas.pringle@southwales.ac.uk (N. Pringle).

http://dx.doi.org/10.1016/j.diin.2014.03.005

happen. Provided the investigator is trained to use these applications as they were intended, the system is inherently assured. The designers consciously choose not to have a write-ability, not because it's just easier that way but because we have a special need to protect the data under investigation.

Assurance standards applicable to digital forensics

Unfortunately there are no explicit rules to define Information Assurance for processing Forensic data. Forensic evidence must adhere to the Daubert principle and the Federal Rules of Evidence in the US, ACPO guidelines in the UK (ACPO, 2012) and corresponding criteria elsewhere. ISO 27037 (ISO 27037:2012, 2012) addresses the acquisition and preservation of digital evidence but uses language such as "protected as far as possible" and that "evidence should be stored in an evidence facility that applies physical security controls". Standards like ISO17025:2005, intended for 'chemical' laboratories, have been the basis of digital forensic facilities but the translation from the analogue to the digital world is not always easy. ISO 27001:2013 defines characteristics of a management system that provides assurance, but not assurance itself. PCI-DSS (PCI Security Standards Council) does provide a more prescriptive standard but doesn't map well to digital forensics. When these are appropriate, unfortunately they are generally based upon the vague notion of 'best practice' and 'the accepted norm' in the particular field. It is difficult to apply in a rapidly developing domain, such as digital forensics, as technology changes are naturally always ahead of 'best practice' developments.

Internal controls in digital forensics

In practical terms, these reveal themselves in some of the characteristics of an existing system when, for example, a new item of evidence is introduced into the lab. It would first be recorded in some form of log. When the evidence image is copied onto the storage facility its success or failure needs to be validated, perhaps with a cryptographic hash digest, for example SHA-1, and this is recorded in the log book. The hash digest is an inherent property of the image. If the validation fails, the operator would investigate the process or equipment and make remedies and rerun the copy. This time, hopefully, it would succeed and the task is complete. Its success, and the previous failure, should both be recorded on the log book. In a paper system, the log book should have certain characteristics. The pages should be numbered and bound together. Anything written should be in ink. Lines on the page should either have writing or be lined through. If the log book is implemented on a computer system there should be an external verification, for example a time date stamp encrypted by PKI, that is beyond the capabilities of the operator to amend. These sorts of controls are common and should be familiar to any investigator.

All these processes should be subject to an *Audit*. By *Auditing*, we are checking that the system worked. The main problem with Auditing is that it is reflective and it often implies a protracted period of time passing before the audit. External audits are often annual, internal audits are perhaps, quarterly. It addresses issues that occurred in the past, assesses their conformance or non-conformance and should trigger changes in the system to prevent further breaches.

This was the case in the quality control employed in most industries in the Western World after the Second World War. Generally, goods were manufactured and were subject to quality control as a final stage where a sample set was tested for conformance. Those non-conforming were removed and either reworked or scrapped. The audit would trigger a period of reflection and perhaps modification to the production system to reduce the failure rate. Regrettably, there was an acceptance that a percentage of non-conformances would get through the system.

From audit to assurance

During the 1960s the Japanese introduced the idea of total quality assurance. The most important aspect of this was that controls were introduced before that action took place, not after.

The dictionary definitions give a sense of the retrospective nature of an audit (Dictionary.com, 2014) and the future intent of Assurance

> ***Audit*** (*noun*)
> 1. **an official examination and verification of accounts and records, especially of financial accounts.**
> 2. **a report or statement reflecting an audit; a final statement of account.**
>
> **Assurance** (*noun*)
>
> 1 **a positive declaration intended to give confidence; a promise**.
>
> synonyms: word of honour, word, guarantee, promise, pledge, vow, avowal, oath, bond, affirmation, undertaking, commitment
>
> 2 **confidence or certainty in one's own abilities.**
>
> synonyms: self-confidence, confidence, self-assurance, belief in oneself, faith in oneself, positiveness, assertiveness, self-possession,
> self-reliance, nerve, poise, aplomb, presence of mind, phlegm, level-headedness, cool-headedness

Japanese production lines did not produce faulty goods because faulty components were not allowed to enter the production line. The effect of this change on the industrial base of the western world is a matter of history. During the 1970s and 1980s products from Japan surged leaving their North American and European competition behind, being viewed as unreliable. Modern management systems like Total Quality Management and Six-Sigma have their focus on controlling inputs and processes during the manufacturing process. Increases in quality, and customer satisfaction, are natural consequences of this approach.

Assurance in current computer systems

Most digital evidence from storage media presented in court is the result of analysis conducted using FTK or EnCase. This is so much a de-facto standard that we rarely question it but both systems are based on the same principles and on more than a decade of acceptance and precedence. At its heart is the idea of always presenting evidence originating 'from the image'.